(12)

NOSC

NOSC / TR 150

AD A047870

NOSC / TR 150

Technical Report 150

RESOLUTION REQUIREMENTS FOR SLOW-SCAN TELEVISION TRANSMISSION OF X-RAYS

A test report for the Remote Medical Diagnosis System

WT Rasmussen, RL Crepeau, and FH Gerber

19 September 1977

Research and Development, November 1976 to January 1977

Prepared for
NAVAL MEDICAL RESEARCH AND
DEVELOPMENT COMMAND

Approved for public release; distribution is unlimited

NAVAL OCEAN SYSTEMS CENTER
SAN DIEGO, CALIFORNIA 92152

NAVAL OCEAN SYSTEMS CENTER, SAN DIEGO, CA 92152

AN ACTIVITY OF THE NAVAL MATERIAL COMMAND

RR GAVAZZI CAPT, USN
Commander

HL BLOOD
Technical Director

ADMINISTRATIVE INFORMATION

Work was performed from November 1976 to January 1977 under sponsorship of the Naval Medical Research and Development Command as part of Program Element 63706N, Project M0096, Task Area M4311 (NOSC S104), Remote Medical Diagnosis System. The test design and protocols were established by WT Rasmussen, PhD (NOSC, Man-System Interaction Division), RL Crepeau (WESTEC Services, Inc), and CDR FH Gerber, MCUSN (Chief of Radiology, Naval Regional Medical Center, San Diego, CA), and the videotape television images were prepared at NOSC by JC Wirthlin (Man-System Interaction Division) and RL Crepeau. The tests were administered to radiologists at the Naval Regional Medical Center, San Diego. Test analyses were performed by WESTEC Services, Inc, San Diego, under Contract N000953-76-M-A286 under the direction of RL Crepeau and WT Rasmussen.

Released by
J Silva, PhD, Head
Man-System Interaction Division

Under authority of
JH Maynard, Head
Command Control-Electronic Warfare
Systems and Technology Department

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER NOSC Technical Report 150 (TR 150)	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER 9
4. TITLE (and Subtitle) 6 RESOLUTION REQUIREMENTS FOR SLOW-SCAN TELEVISION TRANSMISSION OF X-RAYS. A test report for the Remote Medical Diagnosis System.		5. TYPE OF REPORT & PERIOD COVERED Research and Development rept. November 1976 to January 1977
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) 10 W.T. Rasmussen, R.L. Crepeau, and F.H. Gerber		8. CONTRACT OR GRANT NUMBER(s) 16
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Ocean Systems Center San Diego, CA 92152		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 63706N, M0096, M4311 (NOSC S104)
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Medical Research and Development Command National Naval Medical Center Bethesda, MD 20014		12. REPORT DATE 11 19 September 1977
		13. NUMBER OF PAGES 72
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) 14 NOSC/TR-150 12 76p.		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution is unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Medical diagnosis
X-rays
Slow-scan TV transmission of diagnoses
Remote diagnosis

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The Remote Medical Diagnosis System (RMDS) is being developed to link ships and shore stations lacking professional medical staffs with larger ships and stations having such staffs. Although RMDS will be capable of transmitting a variety of visual and auditory medical data (eg, X-rays, ECGs, voice and heart-lung sounds, etc), its most important use will be for the transmission of high-quality X-ray images. Because of system limitations and other design factors, the resolution of the system must be limited to no more than that necessary for functionally satisfactory performance. As used here, functionally satisfactory performance specifically means that the performance of the radiologist in interpreting X-rays and related data transmitted by RMDS must be substantially as effective as his performance in interpreting original X-ray images.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED LB
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

393 159

(20) In order to determine the resolution which will actually be required in use of the system, a series of tests was planned and conducted. In these tests, radiologists at the San Diego Naval Regional Medical Center interpreted test X-rays which were displayed at various levels of resolution. The responses of the radiologists were analyzed to derive the minimum resolution levels which would be compatible with system design and operational requirements. The body of the report describes the experiment, the results obtained, and the conclusions drawn from the results including specific guidelines for resolution needed in the RMDS visual channel.

CONTENTS

CONTENTS (Continued)

ILLUSTRATIONS

TABLES

INTRODUCTION

BACKGROUND

The Naval Ocean Systems Center, under the sponsorship of the Naval Medical Research and Development Command, is performing development work on a medical data transfer network called the Remote Medical Diagnosis System (RMDS). This system will link remote sites (ships or stations) lacking professional medical staffs with diagnostic centers having full medical capability.

Since the RMDS will utilize existing Navy communication systems, it is necessary to stay within the bandwidth constraints imposed by these systems. For this reason only slow-scan television (SSTV) is considered feasible for use by the RMDS to transmit video images. SSTV transmits a static image (ie, does not show motion) and typically may take from 4 seconds to 10 minutes to transmit a single image. The time required to transmit an image is dependent upon the transmission mode (analog or digital), the bandwidth (analog) or data rate (digital), and the level of resolution and gray scale to be achieved. Utilizing a combination of existing Navy communication systems and standard telephone lines, analog transmissions are limited to a bandwidth of 3 kHz and digital transmissions are currently limited to 2400 bits per second (baud). Using a 2400-baud digital link, a system capable of 512 television lines, 256 picture elements per line, and 6 bits per picture element requires 6 minutes and 6 seconds to transmit a single image. Operational demands of Navy communication systems make it desirable to minimize the amount of data to be transmitted in order to constrain transmissions to a reasonable amount of time, and still provide sufficient resolution to allow accurate diagnoses.

Because of these technical and operational factors, the resolution and gray scale to be provided for transferring and displaying radiographic data should be no more than is needed to meet essential requirements. The gray scale requirements for display have been established satisfactorily by existing literature to be at least 6 bits per picture element with 8 bits preferred. No similar guideline applicable to resolution has been established. In order to derive these data on resolution, tests were planned, conducted, and analyzed as described in this report.

The resolution level needed in the RMDS visual channel will be that which will enable the radiologist at the receiving RMDS station to perform satisfactorily his evaluations of the RMDS-transmitted material. This means the radiologist must be able to diagnose advanced pathological conditions and injuries with the RMDS-transmitted material in a manner substantially the same as he would if the original X-ray films were used. It should be noted that the resolution requirement may not be as great as would be the case of making more definitive interpretations typical of normal clinical screening practices.

TEST OBJECTIVE

It is important for technical and cost reasons not to overdesign RMDS with respect to gray scale and resolution. Since data already available establish satisfactorily gray-scale requirements, the tests reported herein are limited to consideration of resolution requirements. Specifically, the objective will be to determine just how much resolution is needed for the radiologist to make correct interpretations of X-ray pictures showing injuries or

advanced pathological conditions. After this resolution level is defined, it will then be possible to design the various components of the system to ensure that the level is maintained operationally on the displays at the receiving station.

TEST APPROACH

In order to determine the level of resolution suitable for RMDS usage, radiologists were asked to evaluate radiographic images presented to them on a video monitor at several levels of resolution. For the tests, 10 radiographs were selected and divided into 2 test groups. Each radiograph was imaged on a special, digital, closed-circuit television system with a defocused camera lens. The degraded electronic images were recorded on video tape in a sequence of images from poorest to best. The tape was then used to present the images on a television screen in the desired test sequence. The presented images were similar to those which would be obtained from an RMDS SSTV system with respect to content, size, orientation, and contrast.

The radiologists were tested individually using the video sequence and then the original X-ray film. Members of the test group 1 were asked to evaluate diagnoses already made with respect to the test items; members of the test group 2 were asked to make their own interpretations of the test items. Neither case histories nor clinical data were available to either group. The diagnoses and evaluations were analyzed subsequently and the responses to the degraded images were compared with the responses to the film readings. Conclusions on resolution requirements were obtained from this analysis.

The remainder of this report is divided into 5 sections. The next section provides a detailed discussion of the equipment and other experimental apparatus and materials. Then, in order, following sections describe the procedures used for the experiment, document the results of the experiment including the analysis of data, discuss the experimental results, and, finally, present conclusions and recommendations derived from the experiment, the latter including specific resolution requirements.

TEST MATERIALS AND EQUIPMENT

GENERAL

Various alternatives were considered for presenting the degraded radiographic images for evaluation. These included a rear-screen video projection system with defocused optics, a closed-circuit television system (CCTV) with various band-limiting filters, and a CCTV with disc- or tape-recorded images. The rear-screen projection system was eliminated because the image presented with such a system was found to be totally noise-free and not representative of a realistic electronic-image output. The CCTV system with tape-recorded images was selected because suitable equipment to implement the other approaches was not available.

RADIOGRAPH SELECTION

It would have been ideal to have large numbers of test images and radiologists for use in the tests in the interest of representative sampling. However, limited personnel resources made this approach impractical. Further, the difficulties anticipated in obtaining time from

radiologists plus the adverse effects of over-long tests, set a limit of one hour on the test period of each radiologist. This meant that only a very limited number of X-ray photographs could be used and that these would have to be selected with special care so as to be a representative sample and provide an adequate testing range. Because pulmonary diseases are the most prevalent of shipboard maladies, half of the films selected for evaluation were chest films. This decision was supported further by the fact that the chest film, because of its physical size and considerable range of detail, was expected to be the most demanding type of radiograph (in terms of information-transfer requirements) to be transmitted.

A detailed evaluation of a number of possible candidate films was performed by a senior radiologist on the study team. This evaluation resulted in a selection of 10 radiographs for use in the test. Appendix A contains photographs of these 10 test films and table 1 lists them along with an *a priori* evaluation of the probable diagnoses. The series included 5 chest films of which 4 were abnormal and one was normal, 2 lateral skull films of which one was normal and one had a fracture, one sinus film, one pelvic film, and one showing a fracture of a major bone. Additional details on the cases and the diagnoses will be provided later on in this report.

TABLE 1. SUMMARY OF TEST FILMS.

Case	Subject (View)	Principal Diagnoses
[illegible]	Bone (femur)	Major fracture with displacement
[illegible]	Chest (AP)*	Atelectasis
		Possible pleural effusion
3	Chest (AP)	Multiple injuries
		Multiple rib fractures
		Fracture of scapula
		Hematoma
		Possible hemothorax
4	Chest (AP)	Normal
5	Chest (AP)	Active tuberculosis
6	Pelvis (PA)**	Fracture (subtle) of right pubic ramus
7	Sinus	Sinusitis
8	Skull (lateral)	Normal
9	Chest (AP)	Pneumothorax
10	Skull (lateral)	Multiple fracture

* Anterior to posterior
** Posterior to anterior

CASE ONE, BONE WITH FRACTURE

This case was that of a femur fracture at approximately mid-length. The fracture showed a displacement of approximately one inch and included one visible bone chip.

CASE TWO, CHEST WITH ATELECTASIS

The radiograph showed a pronounced light area in the lower 20 percent of the lower left lung. This area was clearly abnormal and could have represented either atelectasis or hemothorax, of which the former was deemed correct. No indications resembling pneumonia were present in the film.

CASE THREE, CHEST WITH MASSIVE INJURIES

This radiograph was of a subject who had suffered significant injuries to the right shoulder and lung areas in an auto accident. Specifically, the right scapula had been fractured and displaced, but the fracture itself was not obvious and only the position of the scapula gave the clue. The light region in half of the right lung area indicated a hemotoma or a possible hemothorax condition. The film also clearly indicated an anomaly of the ribs due to their peculiar shape. A very close examination of the ribs indicated that at least 2 were fractured. Diagnosis, however, was made difficult because the film in that area was very poor due to local overexposure. Atelectasis was not a reasonable diagnosis; however, as the film was taken with the patient in a prone position, it was possible to make a misdiagnosis of that condition.

CASE FOUR, CHEST (NORMAL)

This film was of a normal female chest with no visual cues which could have been mistaken for any of the given diagnoses.

CASE FIVE, CHEST WITH ACTIVE TUBERCULOSIS

In this film, the upper lobe of the right lung contained radiologic opacifications, generally beneath rib structures. The shape and structure of these opacifications indicated active tuberculosis but were also suggestive of bacterial pneumonia or pleural plaque. Neither of these latter 2 diagnoses was correct.

CASE SIX, PELVIS WITH FRACTURE OF RIGHT PUBIC RAMUS

The film showed a very subtle fracture of the right pubic ramus. Although there were no clear fracture lines, the indications were slight gradations of gray scale and detail, and the shape of the ramus. There was possibility of misdiagnosis of an intertrochaneric fracture of the left femoral neck due to the fact that a line, representing the edge of the fatty tissue of the buttocks, crossed the femoral neck and gave the appearance of a fracture. This misdiagnosis would be rejected, however, upon careful examination, since this line could be seen to extend outside the femoral neck.

CASE SEVEN, SINUS WITH SINUSITIS

The film for this case showed a significant opacification of the left portion of the face. This could be attributed to either sinusitis or swelling of soft tissue. Due to the lack of a sharp demarcation of the opaque area, sinusitis was concluded to be the most probable diagnosis, barring any other clinical data to support the diagnosis. No true indication of facial fracture appeared; however, in the case of swelling of soft tissue due to a contusion, this could be a logical possibility.

CASE EIGHT, SKULL (NORMAL)

This film of a lateral view of the skull showed a vascular structure running downward from the top central portion of the skull which might appear to be a major fracture under limited resolution conditions. In fact, there was no fracture of any kind. The pineal body on this film was calcified, and thus visible, but not displaced.

CASE NINE, CHEST WITH PNEUMOTHORAX

This film showed a sharp-edged, but subtle pneumothroax of the left lung. No symptoms were present to indicate any of the other listed diagnoses.

CASE TEN, SKULL WITH FRACTURE

This lateral film showed numerous long (up to 5 cm) and distinct fractures located in the center and back of the skull. The pineal body was not calcified, and as such, the pineal displacement diagnosis was meaningless.

EQUIPMENT

The equipment employed in this experiment consisted of the hardware necessary to tape record and subsequently display the selected radiographic images. The main elements were: a television camera, a logarithmic amplifier, a video digitization and analog conversion system, a video tape recorder, and display monitors. Figure 1 illustrates the hardware system configuration used to develop the test tape for the experiment.

TELEVISION CAMERA SYSTEM

Two different camera units were employed to make the test tape. For 8 of the 10 radiographs, a Cohu 7120 camera with a Model 7900 remote controller unit was employed. This is a 32-MHz, 525-line, vidicon television system. The bandwidth was reduced to 8-MHz (−3-dB point) by a plug-in filter in the camera. The measured resolution for this system was 650 horizontal lines per picture height, which agreed with the calculated resolution for these conditions. The Vidicon used in the camera was an RCA, model 8541A, 1-inch, separate-mesh type.

Because the Cohu 7120 was not available when the last 2 radiograph series were made, a minicon camera with essentially the same characteristics was substituted. A Nikkor 28-mm, f/1.8, 35-mm format lens was used with both cameras.

LOGARITHMIC AMPLIFIER

The range of optical densities inherent in radiographs is over 300 to 1, while the range of densities of interest is estimated to be 150 to 1; however, even the latter dynamic range is beyond the capabilities of the television system and video recorder. In particular, if a linear

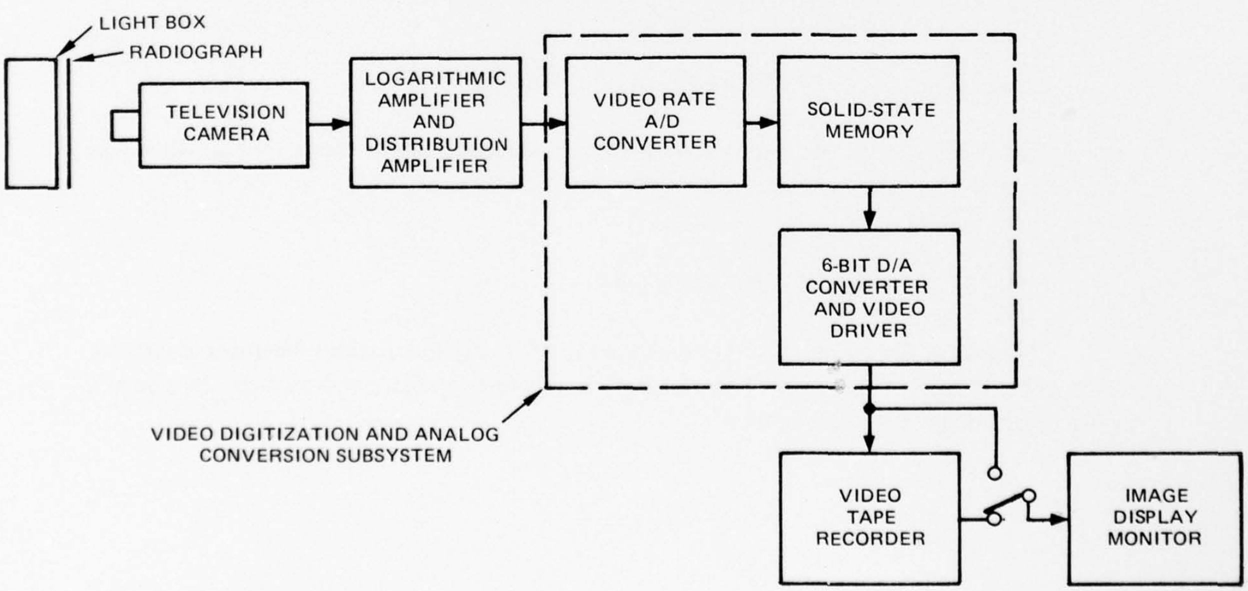

Figure 1. Hardware system for producing test tapes.

response to the film density is employed, much of the image gradation information of interest in the darker portions of the radiograph is lost in the system noise. Since the information in the light portions of the image is usually of less importance, the response in this region can be sacrificed. For this reason, a specially-designed logarithmic amplifier was fabricated and used to compress the dynamic range (response) in the light portions of the radiographs and expand the dynamic range in the dark portions of the radiographs. This amplifier has a bandwidth of 10 MHz and provides response varying from linear to 15 dB of adjustable compression in the light portion.

VIDEO DIGITIZATION AND ANALOG CONVERSION SUBSYSTEM

It was desired to provide images on the test tape to be representative of digitally processed (quantized) images. To do this, an existing real-time video signal digitization, frame storage, and digital-to-analog conversion subsystem was used to produce the test tapes. This system, identified in figure 1, digitizes and stores in real time a frame of data at 6 bits per picture element, reads frame data from memory, and converts these data to quantized (unfiltered) analog form as a 525-line NTSC standard 2-to-1 interlaced-scan signal. The output image contains 975 picture elements in the horizontal direction and forms an image with the commercially standard 4-to-3 aspect ratio.

It should be noted that this approach to simulating the RMDS visual channel necessarily lacks full fidelity, particularly with respect to noise. The differences, as will be explained, made the test situation less demanding than the RMDS with respect to reading through noise. The RMDS will transfer a single image frame to the receiving terminal where it is stored together with all system noise accumulated in transmission and processing. The stored image, including its static noise buildup, is read out of storage 30 times per second for display. In the test system, on the other hand, a new frame was recorded on tape each 1/30 of a second so system noise was randomized from frame to frame on playback. The operational difference is that, in viewing the actual RMDS display, random noise from the transmission and processing becomes an integral part of the stored image whereas, with the test display, the noise forms a random pattern from frame to frame. At the 30-cycle frame ratio, the noise will not summate to a perceptible level for the observer so the system noise which appears randomly on the frame will not be confused with picture noise. On the other hand, in interpreting the RMDS image, the user will see system noise displayed at fixed positions; this will tend to interfere with picture interpretation. (The effect would be similar to over-printing a black and white picture with ink spots.) The magnitude and significance of this difference with respect to noise between the test and actual RMDS displays cannot be determined nor assessed at this time, but conservative interpretation of the experimental findings is advisable for the above reasons.

VIDEO TAPE RECORDER

An IVC 200C, cassette, video-tape recorder was used for recording and playback of the test tape. The resolution of this system for record and subsequent playback was measured at over 550 lines.

DISPLAY MONITOR

The television monitor used to display the test images was a GBC CCTV Corporation, Model MV-17, with a 10-MHz bandwidth and a 17-inch screen. Front-panel brightness and contrast controls were accessible to the radiologists for use during the tests.

RECORDINGS

Because it was desired to limit the participation of each radiologist to one hour, only 10 radiographs at each of 5 levels of resolution were used. This represented a total of 50 images (40 TV and 10 film) to be evaluated by each radiologist. In order to ensure agreement between test-image numbers and the radiologist response sheets, each radiographic image on the test tape was preceded by an image of its respective number. The radiograph images themselves were recorded in such a manner that comparable fields of view and orientations were present for all. For the 4 video-image recordings of each radiograph, the resolution was improved for each succeeding image with the last image at or near the best resolution of which the television-recorder system was capable. The best television recorder system resolution which was measured was about 550 lines, using EIA test charts at the appropriate viewing distance.

In order to determine accurately the degree of resolution being recorded on the tape, this resolution was measured for each of the 40 images. To do this, a bar pattern of known spatial resolution was viewed in place of the film and all of the resolution-influencing factors of the camera configuration were held constant. The camera optics were then defocused (if appropriate) until the bars representing the desired spatial frequency appeared at about half their best contrast ratio.* At this setting, the bars for the next higher resolution were completely defocused while the bars for the next lower resolution were generally sharp. With the bar resolution set, the bar pattern was replaced by the radiograph and the desired video-tape recording was made. Just prior to making each recording, the digital system was tested to ensure that the video signal was optimum; that is, the full dynamic range of the digital system was being used.

Table 2 is a listing of the recordings which were made and the corresponding resolution data. The calculated resolution, in television lines per picture height, is based upon the measured resolution at the film, the measured horizontal field of view, and the 4-to-3 aspect ratio of the television system which was used. Appendix B of this report contains photographs of the recorded images used in the tests.

* This adjustment was performed using the visual display after passing the video through the recorded electronics

TABLE 2. RECORDING PARAMETERS.

Case	Subject	Recording	Resolution (Lp/mm)	Resolution TV Lines	Other
1	Bone fracture	1	0.157	66	F/2.8
		2	0.315	132	Horiz field of view, 11 in.
		3	0.63	264	60-sec recording
		4	1.0	419	
2	Chest with collapsed left lower lobe (AP)	5	0.079	43	F/2.8
		6	0.157	85	Recordings 5, 6, at 90 sec; 7, 8 at 80 sec
		7	0.315	171	Horiz field of view 14.25 in.
		8	0.63	342	
3	Chest with multiple injuries (AP)	9	0.079	50	F/5.6
		10	0.157	99	90-sec recordings
		11	0.315	198	Horiz field of view 16.5 in.
		12	0.63	396	
4	Normal female chest (AP)	13	0.079	47	F/2.8
		14	0.157	94	90-sec recordings
		15	0.315	189	Horiz field of view 15.75 in.
		16	0.63	378	
5	Chest with active tuberculosis (AP)	17	0.079	47	F/4
		18	0.157	94	90-sec recordings
		19	0.315	189	Horiz field of view 15.75 in.
		20	0.63	378	
6	Pelvis with fracture of right pubic ramus (PA)	21	0.157	80	F/2.8
		22	0.315	160	90-sec recordings
		23	0.63	321	Horiz field of view 13.375 in.
		24	1.0	521	Reset log amplifier to compress dynamic range and bring out dark areas while sacrificing bright areas.
7	Sinus with sinusitis	25	0.315	100	F/11
		26	0.42	134	90-sec recording
		27	0.63	201	Horiz field of view 8.375 in.
		28	1.2	303	Reset log amplifier for broader range of gray scales in midrange area.
8	Normal lateral skull	29	0.157	70	F/4
		30	0.315	141	90-sec recordings
		31	0.63	282	Horiz field of view 11.75 in.
		32	1.2	537	
9	Chest (AP) pneumothorax (at expiration)	33	0.079	41	MINICON Camera
		34	1.157	82	F/2.8
		35	0.315	164	90-sec recordings
		36	0.63	327	Horiz field of view 11.625 in. Reset log amplifier for uniformity over dynamic range.
10	Skull (lateral) with multiple fractures	37	0.157	70	MINICON Camera
		38	0.315	141	F/2.8
		39	0.63	282	90-sec recordings
		40	1.2	537	Horiz field of view 11.75 in.

PROCEDURES

TEST CONDITIONS

The tests performed by the radiologists were conducted at the Naval Regional Medical Center, San Diego, in an unoccupied office. The television monitor was placed on a desk with the tape recorder on a nearby table, the radiologist was seated at the desk, approximately 2 feet (60 cm) from the monitor screen, but free to move closer to or further from the monitor for comfortable viewing. The desk space between the monitor and radiologist was available as a working surface for marking the test data sheets. The background light in the room was controlled with venetian blinds to provide good contrast for viewing the monitor.

TEST SUBJECTS

Eight practicing radiologists participated in the tests. They varied in background from a resident with one year of experience to 2 staff members, each with 6 years of experience. The average experience, including residency, was 3.5 years. The 8 radiologists were divided into 2 comparable groups of 4 each for 2 different test approaches. The experience, position, and tests for each radiologist are listed in table 3.

TABLE 3. DATA ON PARTICIPATING RADIOLOGISTS.

			Test	
Radiologist	Position at NRMC	Years of Experience*	(Group 1) Diagnoses Evaluation	(Group 2) Image Diagnosis
A	Resident	2.5	X	
B	Resident	3.0	X	
C	Staff	4.0	X	
D	Staff	6.0	X	
E	Resident	1.0		X
F	Resident	1.3		X
G	Staff	4.0		X
H	Staff	6.0		X

* Includes residency

TESTING SESSIONS

The testing sessions were conducted during normal working hours. A test supervisor was present for the entire test sequence to ensure that the test equipment was ready, to operate the video recorder, to provide the radiologist with the films for viewing, and to answer any procedural questions.

When a radiologist arrived for the test, he was informed that he was participating in a data-gathering experiment and was asked to review a set of instructions concerning the experiment background and procedures. After clearing up any procedural questions, the test supervisor then started the test with the video tapes and continued until the tape was completed and the test data sheets were filled out. (The option of image rerun was always open to the radiologists but was never exercised.)

Upon completion of the video series, the radiologist was then given the original radiographs to evaluate with a conventional view box. These were presented one at a time with sufficient time allowed for evaluation impressions and/or diagnoses. The test was completed when a review of all the radiographs was finished.

APPROACHES

Two basic test approaches were employed. In one, the radiologists (group 1) were given test sheets containing previously prepared diagnoses and were asked to evaluate them. In the second approach, the radiologists (group 2) were asked to supply their own diagnoses. The two test approaches are referred to in this report as the "Diagnoses Evaluation Test" (group 1 test), and the "Image Diagnosis Test" (group 2 test), respectively.

DIAGNOSES EVALUATION TESTS

Of the 8 radiologists tested, every other one, starting with the first, participated in the Diagnoses Evaluation Tests. Appendix C contains a copy of the instructions for those participating in these tests.

Figure 2 is a sample of the test sheet used in the Diagnoses Evaluation Tests. One test sheet was provided for each of the 50 test images which were evaluated. The diagnoses entered on the test sheets were the same for each test image prepared from a given radiograph. The instructions directed each radiologist to evaluate each diagnosis on the test sheet for the test image to which it pertained and to base his evaluation strictly upon what he saw in the test image.

The sets of diagnoses presented in the test series were originally prepared by the radiologist on the study team. The diagnoses included in each set were those which would appear to be reasonable for the respective radiographs and which would require careful interpretation of the radiograph features in order to make a correct selection. Table 4 lists the cases used for the test and the preassigned diagnoses in the order presented for evaluation.

TEST IMAGE # 21 □

IMAGE SUBJECT: PELVIS

COMMENTS:

#1

Diagnosis: FRACTURE OF LEFT PUBIC RAMUS	Definitely Wrong	Probably Wrong	Cannot Decide	Probably Correct	Definitely Correct

#2

Diagnosis: FRACTURE OF RIGHT PUBIC RAMUS	Definitely Wrong	Probably Wrong	Cannot Decide	Probably Correct	Definitely Correct

#3

Diagnosis: INTER-TROCHANERIC FRACTURE OF LEFT FEMORAL NECK	Definitely Wrong	Probably Wrong	Cannot Decide	Probably Correct	Definitely Correct

#4

Diagnosis: NORMAL	Definitely Wrong	Probably Wrong	Cannot Decide	Probably Correct	Definitely Correct

#5

Diagnosis:	Definitely Wrong	Probably Wrong	Cannot Decide	Probably Correct	Definitely Correct

Figure 2. Sample of test sheet used in diagnoses evaluation tests.

TABLE 4. DIAGNOSES PRESENTED FOR EVALUATION IN THE DIAGNOSES EVALUATION TESTS.

Case	Subject	Diagnoses for Evaluation
1	Bone	– Fracture* – Normal
2	Chest (AP)[1]	– Atelectasis* – Hemothorax – Normal – Pneumonia – Pneumothorax
3	Chest (AP)	– Atelectasis – Fracture of scapula* – Hemothorax* – Hematoma* – Rib fracture(s)*
4	Chest (AP)	– Atelectasis – Hemothorax – Normal* – Pneumonia – Pneumothorax
5	Chest (AP)	– Active tuberculosis* – Bacterial pneumonia – Normal – Pleural plaque
6	Pelvis (PA)[2]	– Fracture of left pubic ramus – Fracture of right pubic ramus* – Inter-trochaneric fracture of left femoral neck – Normal
7	Sinus	– Facial fracture – Normal – Sinusitis* – Soft tissue swelling
8	Skull (Lateral)	– Fracture – Normal* – Pineal displacement
9	Chest (AP)	– Atelectasis – Hemothorax – Normal – Pneumonia – Pneumothorax*
10	Skull (Lateral)	– Fracture* – Normal – Pineal displacement

* Correct or potentially correct diagnosis

[1] Anterior to posterior

[2] Posterior to anterior

IMAGE DIAGNOSIS TESTS

Four of the 8 radiologists participating in the study were tested in a manner which required that they provide their own diagnoses based upon what they saw in the test image. These diagnoses were totally unprompted and, further, required that the radiologist indicate his relative confidence in his diagnosis.

Appendix D contains a copy of the pretest instructions provided to the radiologists participating in this second test series. The procedures and test tape for conducting this testing were the same as those for the Diagnoses Evaluation Tests.

Figure 3 shows a sample of the test sheet used in the Image Diagnosis Tests. It may be noted that this sheet provides for up to 4 diagnoses and 5 confidence-level indications ranging from low to high. One sheet was provided for each of the 50 images evaluated.

TEST DATA AND ANALYSIS

PRELIMINARY CONSIDERATIONS

In planning this experiment, it was assumed that, for original photographs and for individual radiologists of roughly equal ability, interpretations for each case would be quite uniform. This did not turn out to be a valid assumption. Consider, for example, Case nine (see table 5). For the diagnosis of atelectasis, Radiologists C and D, the 2 most experienced radiologists in the group, were diametrically opposed in their evaluations. Furthermore, the 4 radiologists were split half-and-half as to whether or not atelectasis was present, while reading the same original film. This lack of agreement occurred in other instances during the testing.

It should be noted that the original films were interpreted last in the testing period, and that the test ran longer than the allocated hour. This means that the film readings may have been rushed in some cases because of the radiologist's need to return to more pressing duties. Also, of course, the readings were made independently and without benefit of supplement patient history and symptomatic information. A review of results obtained in reading the original films is presented later on in this report.

DATA-ANALYSIS APPROACH

The main analysis effort was devoted to the Diagnoses Evaluation Tests. The Image Diagnosis Tests were analyzed to supplement the main analysis effort and assist in drawing conclusions for each radiographic case.

In analyzing the results, the performance of each radiologist working with the television images was compared to his performance when the original films were used. In addition, for each radiologist, the point was determined where, in the television series, the performance was substantially the same as with the original film. This method established the level of resolution needed by each radiologist in order to perform as well with the television images as with the original films. By taking into account the range and distribution of responses, as well as situational factors for each case, the analysts derived subjectively resolution levels which appeared to be appropriate for the radiologists as a group.

IMAGE SUBJECT: PELVIS

COMMENTS:

#1

Diagnosis:	CONFIDENCE				
	LOW	MEDIUM-LOW	MEDIUM	MEDIUM-HIGH	HIGH

#2

Diagnosis:	CONFIDENCE				
	LOW	MEDIUM-LOW	MEDIUM	MEDIUM-HIGH	HIGH

#3

Diagnosis:	CONFIDENCE				
	LOW	MEDIUM-LOW	MEDIUM	MEDIUM-HIGH	HIGH

#4

Diagnosis:	CONFIDENCE				
	LOW	MEDIUM-LOW	MEDIUM	MEDIUM-HIGH	HIGH

Figure 3. Sample of test sheet used in image diagnosis tests.

TABLE 5. CASE NINE, CHEST WITH PNEUMOTHORAX

Test Images 33 through 36 and 49

Radiologist	Resolution		Atelectasis	Hemothorax	Normal	Pneumonia	Pneumothorax
	Lp/mm	TV Lines					
A	0.079	41	0	–	0	0	0
	0.157	82	–	–	0	+	–
	0.315	164	––	––	––	––	++
	0.63	327	––	––	––	––	++
	Radiograph		––	––	––	––	++
B	0.079	41	0	+	––	0	0
	0.157	82	–	+	–	0	0
	0.315	164	+	0	––	–	++
	0.63	327	+	0	–	–	++
	Radiograph		+	––	––	––	++
C	0.079	41	–	+	––	+	–
	0.157	82	–	+	––	+	–
	0.315	164	0	0	––	0	+
	0.63	327	++	0	––	––	++
	Radiograph		++	––	––	––	++
D	0.079	41	0	0	0	0	0
	0.157	82	0	0	0	0	0
	0.315	164	–	–	–	–	–
	0.63	327	–	–	–	–	–
	Radiograph		––	––	––	––	––

++ Definitely Correct
+ Probably Correct
0 Cannot Decide
– Probably Wrong
–– Definitely Wrong

QUALITATIVE CATEGORIZATION ANALYSIS

As was shown in figure 2, the test sheets used in the Diagnoses Evaluation Tests used 5 categories for responses to each preassigned diagnosis. These were, definitely wrong, probably wrong, cannot decide, probably correct, and definitely correct. In the qualitative rating analysis, the evaluation made by each radiologist (group 1) using these 5 categories on each preassigned diagnosis with the television series was compared to the evaluation made by the same radiologist working with the original film. The performance of each radiologist on each case was then categorized using the data for all of the preassigned diagnoses for the particular case.

COMPLETE AGREEMENT CATEGORY

This first category of performance represented the case in which the individual diagnosis evaluations agreed completely, on a one-for-one basis, with the original film-reading results. For example, in table 5, the evaluations of radiologist A at resolutions of 0.315 lp/mm and 0.63 lp/mm would be placed in this category because there is 100 percent agreement with the film-reading results.

MINOR DISAGREEMENT CATEGORY

This category was used when the following conditions were met in comparing the evaluations of the diagnoses: (a) all recorded television-based evaluations agreed "in direction" with the film-reading results; and (b) no "cannot decide" evaluation of any diagnosis was given.* By this definition, the results of radiologist D for 0.315 lp/mm and 0.63 lp/mm image evaluations of table 5 would be categorized as "minor disagreement." The results of radiologist C for the same resolution would also be categorized as "minor disagreement" because of the "cannot decide" evaluation of the hemothorax diagnosis.

MODERATE DISAGREEMENT CATEGORY

This category was used when either of the following conditions were met when comparing the diagnoses evaluations: (a) the evaluation of no more than one diagnosis in a direction opposite to the evaluation made from film reading but erring conservatively; and (b) one "cannot decide" evaluation was given.

"Erring conservatively" as used here is defined as evaluating an "abnormal" diagnosis as probably or definitely correct when the subsequently film-based evaluation resulted in a "probably" or "definitely wrong" evaluation for that diagnosis. Clearly, this approach has shortcomings. In a real-life situation, the diagnosis of a condition as existing when in fact it did not exist could result in a prescription of treatment which might be detrimental to the patient's true condition. The rationale of this approach, however, is that a misdiagnosis of a condition as being existent would likely result in a more extensive investigation of the patient's condition. Thus, additional examination and subsequently derived information should correct the initial erroneous diagnosis.

---

* If, for a diagnosis, the film-reading evaluation was "definitely correct," a "probably correct" evaluation of the video image would agree in direction. Furthermore, "probably wrong" or "definitely wrong" evaluation would be considered to be in the opposite direction to the film-reading result.

Based upon these criteria, the diagnoses evaluations by radiologist C at 0.63 lp/mm in table 5 were in the "moderate disagreement" category because of the "cannot decide" evaluation for the hemothorax diagnosis.

SERIOUS DISAGREEMENT CATEGORY

This category was used when the diagnoses evaluations were worse than the "moderate disagreement" category and, specifically, when any of the following conditions were found: (a) evaluation of more than one diagnosis was in the direction opposite to that of the film reading; (b) the recorded television-based diagnosis was such as to evaluate one or more abnormal conditions as normal; and (c) when more than one "cannot decide" evaluation was given.

MAJORITY VOTE ANALYSIS

A routine procedure in the evaluation of radiographs is to employ a group of radiologists in the film-reading process. The resulting diagnosis represents a consensus of opinion based upon the combined knowledge and the interchange of viewpoints among the radiologists. In using this procedure, the combined knowledge, experience, and perceptions of the group yield a more accurate diagnosis than can be provided, on the average, by a single radiologist.

In the tests reported upon in this document, however, the radiologists worked independently. It was thought to be of interest to determine what test results might have been achieved had the radiologists worked as a team. In order to obtain some measure of the team-approach effectiveness, a majority vote analysis approach was used to analyze selected data. The procedure was simply to determine the majority opionion of the 4 group 1 radiologist's evaluations for each case, diagnosis, and resolution combination.

The rules for the decision-making process of the majority vote analysis were: (a) of the 4 given evaluations if any one is in the majority, select that evaluation; (b) if the 4 evaluations are all different, assign an evaluation of "cannot decide"; and (c) if the 4 separate evaluations consist of 2 evaluations only, then the majority vote evaluation would consist of both listed evaluations.

This majority approach is very simple and, as such, probably has shortcomings as a model for the actual staff-team analysis situation. On the other hand, alternative schemes such as weighting the votes by radiologist experience cannot be assured as being more accurate. This is for the reason that other factors such as personality, specialties, and the like would also influence the group result and would be too complicated to be modeled in this present effort. Thus, the selected simple majority vote approach makes the fewest assumptions and is likely to be as accurate as any model within the scope of this study. The results of the majority vote analysis for the Diagnoses Evaluation Test are included in Appendix H.

TEST RESULTS

The results of the test are presented in 4 subsections: radiographs by all cases; radiographs for chest cases only; television images by cases; and television images by diagnosis. In the following paragraphs, the results of the radiologists examination of the original radiographs for each case are discussed. No discussion is presented of the bone fracture case since

the gross nature of the subject material did not provide any significant data other than the fact that all radiologists recognized the fracture.

CASE TWO, CHEST WITH ATELECTASIS

In the Diagnosis Evaluation Test, all radiologists in group 1 evaluated the "normal" and "pneumothorax" diagnoses as "definitely wrong" and the "atelectasis" diagnosis as "definitely correct." Three of the 4 radiologists evaluated "hemothorax" as "probably correct" while the fourth could not decide on that diagnosis. These evaluations reflect the fact that both atelectasis and hemothorax present basically the same appearance, a radiographically opaque area. Resolving the conflict of diagnosis between atelectasis and hemothorax would have to make use of more careful film examinations or, more probably, patient history or symptomatic data.

For the Image Diagnosis Tests, all 4 radiologists in group 2 indicated a possible pleural effusion; fluid in the pleural cavity. Two of the radiologists considered atelectasis a possibility.

For the Diagnoses Evaluation Tests, 2 of the 4 radiologists in group 1 believed that "pneumonia" was probably correct. In the Film Diagnosis Tests, only one of the 4 group 2 radiologists indicated a possibility of lower left-lobe (LLL) pneumonia.

From an analysis of the majority votes of the group 1 radiologists in the Diagnoses Evaluation Tests, the following results were obtained: Atelectasis, "definitely correct"; Hemothorax, "probably correct"; Normal, "definitely wrong"; and Pneumothorax, "definitely wrong."

In the case of pneumonia, a split decision between "probably correct" and "probably wrong" resulted effectively in a "cannot decide" evaluation. The split decision, together with the fact that one radiologist in group 2 made a diagnosis of "LLL pneumonia," indicates that some confusion existed as to the pathological nature of the lower left lung opacification. This confusion probably could not have been resolved from the film image and thus could not have been resolved from the television images.

CASE THREE, CHEST WITH MASSIVE INJURIES

Pre-experiment analysis by the project radiologist indicated diagnoses for this case of multiple rib fractures, fracture of the right scapula, and hematoma. Hemothorax was considered a possible present condition. In the Diagnoses Evaluation Tests, all radiologists in group 1 agreed with these evaluations except that radiologist C could not decide the existence of scapula fracture. The evaluation responses to the atelectasis diagnosis were mixed and varied from "probably correct" to "probably wrong" and included a "cannot decide." The fact that there were massive chest injuries probably made this latter diagnosis appear reasonable. It should be noted that the participating radiologists were instructed to base their diagnostic evaluations only upon what they saw in the pictures. It is not known, however, whether the atelectasis evaluations were based upon direct visual presentations or on indirect consideration of the fact that an obvious pulmonary contusion condition was present.

A summary of the diagnoses by radiologists in group 2, the Image Diagnosis for this case, is presented in table 6. Note that the rib fracture and hemothorax (ie, pleural effusion/fluid) conditions were correctly diagnosed by all radiologists. The scapula fracture was seen by only 2 of the 4 participants, and 2 radiologists diagnosed a lesion in the right lung.

TABLE 6. DIAGNOSES BY GROUP TWO RADIOLOGISTS FOR CASE THREE.

DIAGNOSIS	CONFIDENCE EVALUATION BY RADIOLOGIST			
	E	F	G	H
Multiple rib fracture	H	H	H	H
Pleural effusion/fluid	H	H	M-H	H
Pleural mass/lesion	–	H	H-M	–
Fracture of scapula	H	–	–	H

H = high M = medium

CASE FOUR, NORMAL CHEST

In the Diagnoses Evaluation Tests, all radiologists in group one agreed that the film was probably or definitely normal. The abnormal diagnoses presented were all evaluated as "definitely wrong" by all radiologists.

In the Image Diagnosis Tests, the evaluations of the radiologists in group two were: "normal" (radiologist E); "mass left hilum" (radiologist F); "centrally enlarged pulmonary arteries (radiologist G); and "left hilar density, probably large pulmonary artery segment" (radiologist H). In the absence of obvious symptomatic conditions, radiologists F through H presented the most apparent abnormality they could find. The radiologists generally concurred in what they saw as the most apparent abnormality in this normal radiograph.

CASE FIVE, CHEST WITH ACTIVE TUBERCULOSIS

In the Diagnoses Evaluation Tests, all radiologists in group one agreed the film was not normal. The only inconsistencies in the evaluation took place in the bacterial pneumonia diagnosis. For this diagnosis, 2 radiologists could not decide on the evaluation, and the remainder evaluated the diagnosis as "probably wrong" and "probably correct." With regard to the pleural plaque diagnosis, other than the two "cannot decide" evaluations, the other radiologists indicated that this diagnosis was "probably wrong." Three of the 4 radiologists believed that active tuberculosis was "probably correct" and one could not decide.

When a majority vote analysis was conducted, the results were: active tuberculosis, "probably correct"; bacterial pneumonia, "cannot decide"; normal, "definitely wrong"; and pleural plaque, "cannot decide." Clearly, the original film did not provide sufficient information to definitely diagnose the presence of bacterial pneumonia or pleural plaque.

In the Image Diagnosis Tests, all of the radiologists in group 2 detected the 2 right upper lung nodules; 3 of the radiologists referred either to tuberculosis or the more general diagnosis of granulomas disease. Other diagnoses referred to lower/middle left-lung densities. The fact that, in this test, the radiologists did not specifically mention pneumonia or pleural plaque mitigates against any detrimental conclusions concerning the uncertainty in the

Diagnoses Evaluation Test results for the pneumonia and pleural plaque diagnoses by group 1. The presumably correct diagnosis of active tuberculosis was evident. The uncertainty in the pneumonia and pleural plaque diagnoses may have come about by the radiologists' attempts to find very early, and thus, subtle signs of these diseases. The problem could have been avoided if the diagnoses had been given as "advanced bacterial pneumonia" and "extensive pleural plaque" instead as presented.

CASE SIX, PELVIS WITH FRACTURE

The Diagnoses Evaluation Test results indicate that all radiologists in group 1 agreed that the right pubic ramus was fractured, while the left was not. One difference of opinion occurred concerning the inter-trochaneric fracture of the left femoral neck. Radiologist C failed to discern the fat stripe from a fracture line and thus wrongly evaluated this diagnosis as "definitely correct."

The majority vote analysis results show that the correct diagnosis of a fracture of the right pubic ramus would be evaluated as "definitely correct." The remaining diagnoses would be evaluated as "definitely" or "probably wrong."

In the Image Diagnosis Test results, 3 of the 4 radiologists in group 2 indicated a diagnosis of fracture of the right ischium and pubic ramus. These diagnoses are consistent with the pre-experimental analysis of this case film. One radiologist did not see the fracture, and instead, listed a diagnosis of a right pelvic mass or lytic lesion.

CASE SEVEN, SINUS WITH SINUSITIS

In the Diagnoses Evaluation Test results, all radiologists in group 1 agreed the film subject was definitely not normal and that the sinusitis diagnosis was either "probably" or "definitely correct," and that facial fracture was a wrong diagnosis. In the case of soft tissue swelling, the radiologists were split, with one half evaluating this diagnosis as correct and the other half evaluating the diagnosis as wrong. Thus, the majority vote analysis yields a "cannot decide" for the soft tissue swelling diagnosis.

In the Image Diagnosis Tests, all 4 radiologists in group 2 indicated some form of fluid in the left sinus, with 2 radiologists specifically indicating sinusitis. Two of the radiologists also indicated a small air fluid level in the left maxillary sinus. One radiologist also indicated a diagnosis of a fracture of the R/O orbital floor.

The soft-tissue swelling diagnosis was not uniformly evaluated by the radiologists based upon their film readings. Had patient-condition data concerning a possible facial contusion been available, the evaluation of this diagnosis would undoubtedly have been uniformly the same.

CASE EIGHT, NORMAL SKULL

The Diagnoses Evaluation Tests for this case produced inconsistent results in the film reading. The fracture diagnosis was evaluated as wrong by 3 radiologists in group 1 and "definitely correct" by the fourth. No apparent fracture existed on the film, and the diagnosis of the fracture may have been a misdiagnosis of a supposed hairline fracture. In so far as the

pineal displacement diagnosis was concerned, one radiologist evaluated this as "definitely wrong" while the remainder indicated a "cannot decide" evaluation. These later evaluations may have been due to the lack of a measurement instrument, as the pineal was clearly visible. The majority vote analysis results indicate a fracture as "probably wrong" and the normal and pineal displacement diagnoses evaluations as "cannot decide."

In the Image Diagnosis Tests, 2 of the 4 radiologists in group 2 evaluated the film as normal. The remaining 2 noted carotid calcification, with one radiologist further noting dorsum demineralization. While those are correct diagnoses, neither is considered significant to the test.

CASE NINE, CHEST WITH PNEUMOTHORAX

In the Diagnoses Evaluation Tests, the radiologists in group 1 agreed perfectly on their evaluations of all of the diagnoses, except atelectasis. In the case of atelectasis, 2 radiologists evaluated this diagnosis as wrong, while the remaining two evaluated this diagnosis as correct.

Any symptoms of atelectasis must have been extremely subtle. The lack of comments by the radiologists provided no indication of what prompted a positive evaluation for this diagnosis. One radiologist who evaluated atelectasis as "definitely wrong" provided the following comment: "There is also a pneumomediastinum which did not transmit at all."

Using the majority vote analysis for group 1, all diagnoses presented, except the correct one, would be evaluated as "definitely wrong." The correct diagnosis would be evaluated as "definitely correct." In the Image Diagnosis Tests, all radiologists in group 2 diagnosed the film as a case of pneumothorax.

CASE TEN, SKULL WITH FRACTURE

In the Diagnoses Evaluation Tests, all radiologists in group 1 agreed a fracture definitely existed. In that the pineal was not calcified, no decision was made on the pineal displacement diagnosis. The Image Diagnosis Test results consisted of all 4 group 2 radiologists diagnosing multiple fractures.

COMPARISON OF FILM-READING RESULTS FOR CHEST CASES ONLY

NORMAL DIAGNOSES FOR CHEST CASES

Consider the results of the Diagnoses Evaluation Tests with respect to the preassigned normal diagnoses (Appendix F). In this case, the abnormal films (cases 2, 5, and 9) are all correctly evaluated as such; the normal diagnosis is evaluated as "definitely wrong." In case four, in which there was no abnormality, all radiologists indicated the film as probably or definitely normal.

HEMOTHORAX DIAGNOSES FOR CHEST CASES

In the 4 chest cases where hemothorax was a diagnosis for evaluation, the evaluation results from the film were generally consistent among the radiologists. The cases where hemothorax was not the correct diagnosis (4 and 9), the diagnoses were consistently evaluated as "definitely wrong." In case three, where hemothorax was possibly correct, all radiologists evaluated the diagnosis as "probably correct." In case two, where the actual diagnosis was atelectasis, 3 of the 4 radiologists felt that pneumothorax, which is visually similar to atelectasis, was "probably correct."

PNEUMOTHORAX DIAGNOSES FOR CHEST CASES

Three cases included the diagnoses of pneumothorax (2, 4, and 9). All radiologists in all 3 cases were able to accurately evaluate the diagnosis.

ATELECTASIS DIAGNOSIS FOR CHEST CASES

Four cases included the diagnosis of atelectasis (2, 3, 4, and 9), with one actual case (2) of atelectasis. The radiologists were in agreement and accurate in their evaluation of the atelectasis diagnosis for case 2 and in the normal chest film, case 4. In the chest cases with pneumothorax (9) and the case with massive injuries (3), the radiologists disagreed in their diagnoses. In these cases, the majority vote analysis results were: case three – Atelectasis, "definitely correct"; case nine – Atelectasis, "definitely wrong." Thus, with the majority vote analysis, the only discrepancy was in Case three where massive injuries existed and atelectasis was not considered a correct diagnosis.

PNEUMONIA DIAGNOSES FOR CHEST CASES

The pneumonia diagnosis was evaluated in 4 cases (2, 4, 5, and 9) and was not the correct diagnosis in any of these cases. In the cases of the normal chest film (4) and the pneumothorax case (9) all radiologists indicated that pneumonia was "definitely wrong." In cases where active tuberculosis (5) and atelectasis (2) were the actual diagnoses, the radiologists provided conflicting viewpoints in their evaluations. From these latter results, it seems evident that without further supporting information, visual symptoms of bacterial pneumonia as typified by cases two and five are not amenable to accurate radiographic interpretation.

TV-BASED IMAGE RESULTS BY CASES

The following paragraphs discuss the results of the radiologists' examination of the television images during the testing. Utilizing the qualitative rating approach discussed earlier, the data from the Diagnoses Evaluation Tests of group 1 were analyzed. Appendix H presents the results for each level of resolution, including the majority vote data. (Because of the straightforward results obtained for the bone fracture (case one) it is not summarized in Appendix H.)

CASE ONE, BONE FRACTURE

Even at the lowest resolution of 0.157 lp/mm, all radiologists were able to evaluate and diagnose the subject image as a fracture. At the third level of resolution, 0.63 lp/mm, one group 3 radiologist in the Image Diagnosis Test indicated that the fracture involved a small fragment between the major fragments. This bone fragment measured about 6.5 mm by 3.3 mm and was separated from the main fragment by about 0.8 mm along a curved line the length of the chip. This length lay in the vertical direction of the television image.

System resolution good enough to allow recognition of bone fragments should also be good enough for positive recognition of fractures. Thus, fracture lines with widths on the order of the system resolution (in line pairs per millimeter) should be readily detectable in films transmitted with the RMDS system. This statement must be qualified by the assumptions that the fracture is shown roughly in the vertical direction on the television monitor, and that the length of the fracture is greater than about 4 times the fracture width.

CASE TWO, CHEST WITH ATELECTASIS

At the lowest resolution of 0.079 lp/mm, all of the radiologists were able to determine that the film was not normal. Furthermore, based upon the comments from the Image Diagnosis Test at this resolution, all radiologists in group 2 felt that effusion in the lower left pleural was possible. (Confidences ranged from low to medium.)

As the resolution was improved, all radiologists in group 1 (the Diagnoses Evaluation Test) concluded that the atelectasis diagnosis was "probably" or "definitely correct." At the second level of resolution, 0.157 lp/mm, in the Image Diagnosis Tests, 3 of the 4 radiologists in group 2 had indicated atelectasis as a possibiliy, with a medium to medium-high confidence level.

In both the Diagnoses Evaluation Test and the Image Diagnosis Test, the pneumothorax diagnosis was either discounted or not mentioned. Furthermore, in both tests, the evaluations of pneumonia and hemothorax versus atelectasis were not unanimously resolved in either the films or the television-based images.

The majority vote results for the Diagnoses Evaluation Tests indicate that all of the resolution levels produced some results in the "serious disagreement" category. On the other hand, in the Image Diagnosis Test, results at the third resolution level, 0.315 lp/mm, 3 of the 4 group 2 radiologists considered atelectasis a possibility with medium to high confidence. The next most agreed-upon diagnosis was pleural scarring (by 2 of the 4 radiologists). Also it must be noted that all 8 radiologists determined this case was not normal even at the lowest resolution level.

Based upon the test results, but weighing the results of the Image Diagnosis Test more heavily, the 0.315-lp/mm resolution would appear to be satisfactory for analysis of this case. Thus, at this resolution, the results will be as good as can be achieved under the test circumstances.

CASE THREE, CHEST WITH MASSIVE INJURIES

In the Diagnoses Evaluation Test, and at the third resolution level, 0.315 lp/mm, 2 of the 4 radiologists in group 1 indicated a "cannot decide" on two diagnoses and reversed their opinion on a third diagnosis, when reading the original radiograph. A third radiologist

indicated a "cannot decide" evaluation on 4 of the 5 diagnoses. The fourth radiologist had nearly perfect agreement with his radiograph reading results.

In the Image Diagnosis Tests, it was at 0.315-lp/mm resolution that 3 of the radiologists in group 2 defined the EKG pad on the left chest as other than a nodule in the left lung. One radiologist continued to evaluate this EKG pad as a nodule even at the best television resolution (0.63 lp/mm). At 0.315-lp/mm resolution, the radiologists' comments appear to indicate that the patient's recumbent position for the radiograph was not obvious. The obviously misshaped right lung was apparent but its source was not defined.

At 0.63-lp/mm resolution, the qualitative evaluations of the group 1 radiologists evaluations were divided between "moderate disagreement" and "minor disagreement," as given in table 2. For 0.63-lp/mm resolution, and for the Image Diagnosis Test, half (2) of the radiologists in group 2 indicated hematoma (in equivalent terms), while all indicated pleural effusion of the right lung. Further, 2 of the group 2 radiologists indicated fracture of the right ribs. In these tests, and at this resolution, none of the group 2 radiologists noted the fracture of the scapula, although when reading the original film, 2 did note this diagnosis. However, at this resolution, all 4 group 1 radiologists indicated "definitely correct" for the fracture of the scapula, and 3 of the 4 indicated "probably correct" at the 0.315-lp/mm resolution.

The information content of this drastic chest trauma case was found to be too great for the best television resolution and analysis conditions employed. Revisions of either of these factors, eg, longer and more detailed examination, could possibly make images at the (best) resolution useful. Clearly, any poorer resolution would be inadequate, although since group 1 radiologists were able to diagnose the fractured scapula, this would appear to indicate that given some additional information (such as clinical or symptomatic data), adequate diagnoses could be made even at a slightly lower resolution.

CASE FOUR, NORMAL CHEST

For the Diagnoses Evaluation Tests, and at the second resolution level, 0.157 lp/mm, 3 of the 4 group 1 radiologists indicated "cannot decide" evaluations for 2 or more of the 5 diagnoses. The fourth radiologist's evaluations were in the "minor disagreement" category. At the third level, 0.315 lp/mm, only 2 of the radiologists made a "cannot decide" evaluation, and each was for the pneumothorax diagnosis. The overall quality of the evaluations at this resolution, based upon majority vote analysis, was between the "moderate disagreement" and "minor disagreement" categories, but somewhat closer to the latter.

In the Image Diagnosis Tests, one group 2 radiologist evaluated all resolutions as normal. The remaining 3 radiologists established diagnoses which, at 0.157-lp/mm resolution, related to hilar mass/enlarged pulmonary arteries. They retained these diagnoses throughout the remainder of the resolutions and the film reading with increasing confidence levels. The left pulmonary artery was enlarged, but this was neither *a priori* nor *a posteriori*, considered a significant pathological condition.

The results of this case tend to indicate that 0.315 lp/mm of resolution is satisfactory for establishing the normality of the chest radiograph vis-a-vis moderately advanced pathologic conditions other than pneumothorax.

CASE FIVE, CHEST WITH ACTIVE TUBERCULOSIS

Effective analysis of the results from the Diagnoses Evaluation Test for this case were hampered by the fact that 2 of the group 1 radiologists evaluated 2 and 3 diagnoses as "cannot decide" upon reading the original radiograph. In fact, Radiologist D evaluated all of the diagnoses, except normal, as "cannot decide" for all resolutions. At the third level, 0.315 lp/mm, this radiologist indicated that the normal diagnosis was "definitely wrong." The 2 radiologists who did provide decisive evaluations for the Diagnoses Evaluation Tests both performed with "complete agreement" at the fourth level, 0.63-lp/mm resolution, but with "moderate disagreement" and "minor disagreement" at the third level (0.315 lp/mm).

In the Image Diagnosis Tests, 2 radiologists in group 2 had located the true pathology (upper right-lung nodules) at the lowest resolution (0.079 lp/mm). In one case, the nodules, which lay under adjacent ribs, were defined as "old rib fractures." The confidences for these diagnoses were "low." At the second level, all radiologists had detected the upper right-lung nodules, but no definitive diagnoses, other than "rib fractures," were given. At the third level, one radiologist suggested "probably TB" with a medium-high confidence level. At the highest level, 3 of the 4 radiologists indicated diagnoses of tuberculosis, or some equivalent variant thereof. The confidences in these diagnoses were medium and high.

The results of this case indicate that the detection of tuberculosis of this nature requires between 0.315 and 0.63 lp/mm of resolution for whole chest views. A system with an overall resolution as low as 0.157 lp/mm would be practical for this case, assuming that a zoom capability existed. Such zoom capability would have to allow for a linear magnification of at least 4 to 1.

CASE SIX, PELVIS WITH FRACTURE OF RIGHT PUBIC RAMUS

It was recognized *a priori* that, due to the subtle nature of the symptom of this case, increased resolution over the chest film cases would be required. As such, the range of resolutions used was 0.157 to 1.0 lp/mm.

For the Diagnoses Evaluation Tests, and at the second level, 0.315 lp/mm, 75 percent of the diagnosis evaluations by group 1 were "cannot decide," indicating the unacceptability of this resolution level. At the third level, 0.63 lp/mm, 3 of the 4 radiologists still indicated one "cannot decide" evaluation among those diagnoses other than normal. Nonetheless, one radiologist did provide a perfect agreement with his radiograph reading results which were very accurate. At the best television resolution of 1.0 lp/mm, 2 radiologists' evaluations were in the "minor disagreement" category, while the other 2 were in the "moderate" and "serious disagreement" categories. The majority vote analysis at the highest resolution level, 1.0 lp/mm, resulted in a categorization of "minor disagreement."

Three of the 4 group 2 radiologists in the Image Diagnosis Test were able to detect the right pubic ramus fracture in the original film. At the best television resolution, only one radiologist mentioned a pelvic fracture. The confidence level on this diagnosis was only medium. The remaining television image-based diagnoses (at this resolution) related to the lesions, the sacroiliac, and the bone mineralization pathology.

The negative results of the Image Diagnosis Test tend to mitigate strongly against the relatively good results of the Diagnoses Evaluation Test at highest resolution. This stems from the fact that the Diagnoses Evaluation Test prompted the viewer to consider certain possibilities. In the more realistic case of the Image Diagnosis Test, the viewer was required to locate

the problem, assisted neither by diagnostic cues nor by symptomatic data. Thus, less than 1.01p/mm of resolution in this case would very likely be unacceptable. A 1.0-1p/mm resolution would probably be only a marginally acceptable one, although, again, higher resolution could be attained with a zoom capability given that the radiologist knew where to look for abnormal condition.

CASE SEVEN, SINUS WITH SINUSITIS

Due to the small size of the radiograph, the test resolutions ranged from 0.315 to 1.2 lp/mm. Its visual symptoms were *a priori* judged to be mostly related to the gray-level change attributable to the opacification associated with the sinusitis condition. Consistent with this assumption is the fact that, for the Diagnoses Evaluation Test and a resolution of 0.63 lp/mm, each group 1 radiologist evaluated one of the abnormal diagnoses as "cannot decide." Furthermore, 3 of the 4 radiologists evaluated the facial fracture diagnosis as "probably correct," both at this resolution and at the best television resolution, while evaluating it as wrong upon reading the original radiograph. At the best television resolution, the quality of the evaluations for the 4 radiologists were categorized as follows: two "serious disagreement," one "moderate disagreement" and one "minor disagreement."

In the Image Diagnosis Tests, all group 2 radiologists were able to recognize the opacification of the left sinus area at the lowest resolution of 0.315 lp/mm. At this resolution, 2 of the 4 radiologists identified sinusitis, or the roughly equivalent diagnosis of sinus mucosal thickening, with medium-low confidence levels. At the second level, 0.42 lp/mm, 3 radiologists had presented one of these preceding diagnoses. Also at this resolution level, 2 radiologists presented the diagnosis of a fracture with low confidence. At the third level, 0.63 lp/mm, the radiologists' evaluations were about the same as at the second level, except that 3 of the 4 radiologists evaluated a fracture as a possibility with confidences ranging from low to high. At the highest resolution of 1.2 lp/mm, 2 radiologists still indicated a fracture diagnosis, while 3 diagnosed either sinusitis or mucosal thickening.

While the film showed no fracture, a dark streak in the medial wall of the left orbit was visible. This streak probably represents an air-fluid differential, and was mistaken for a fracture when reviewing the television images. When the film itself was evaluated, none of the 8 radiologists indicated fracture diagnoses.

This case again appears to generate conflicting results between the Diagnoses Evaluation Test and the Image Diagnosis Test. In an actual clinical situation, patient history and physical symptomatic condition information would be available. With this information, and at the second resolution level, 0.42 lp/mm, a radiologist could easily diagnose whether or not a fracture existed and could also differentiate between sinusitis and soft tissue swelling. Thus, this resolution level would be adequate for this type of case.

CASE EIGHT, NORMAL SKULL

The tests with this normal skull film were to determine two things: first, could the large vascular structure (approximately 2.5 mm wide and over 15 cm long) be discerned as *not* being a fracture; and second, could the calcified pineal (of about 5-mm diameter) be identified.

In the Diagnoses Evaluation Test, and at the second resolution level, 0.315 lp/mm, 3 of the 4 group 1 radiologists indicated a decisive diagnosis of the fracture. Two indicated a fracture existed and one indicated that it was not a fracture. At the third and fourth levels (0.63 lp/mm and 1.2 lp/mm), all the radiologists had an opinion; they were evenly split as to whether or not the visual symptom represented a fracture. In viewing the actual film, 3 of the 4 radiologists changed their opinions.

In the Image Diagnosis Test, at the second level, 0.315 lp/mm, one group 2 radiologist diagnosed the prominent vascular structure as a fracture with low confidence. At the third level, 0.63 lp/mm, 3 of the 4 radiologists mentioned this visual sympton; 2 recognized it as a large vascular groove and one diagnosed it as a fracture. All of these diagnoses were with medium-low to medium confidence. For the best resolution of 1.2 lp/mm, 2 of the radiologists indicated the equivalent of normal diagnoses, while a third felt, with low confidence, that there was a fracture. The fourth radiologist could not decide between a fracture and a prominent vascular groove.

From the results, it is clear that 0.315 lp/mm of resolution made the pineal visible and was suitable for displacement measurement, so a resolution of twice this, or 0.63 lp/mm, is considered adequate for this type of image.

CASE NINE, CHEST WITH PNEUMOTHORAX

In the Diagnoses Evaluation Test at the lowest resolution level, 0.079 lp/mm, 3 of the group 1 radiologists supplied "cannot decide" evaluations of the diagnoses. The fourth radiologist supplied evaluations for all 4 of the non-normal diagnoses, but all were wrong evaluations. At the second level, 0.157 lp/mm, the diagnostic evaluations provided were generally either in error or "cannot decide." At the third level, 0.315 lp/mm, all but one radiologist had diagnoses evaluations which were categorized as either "moderate disagreement" or "minor disagreement," but in quality from poor to good; all 4 correctly diagnosed the true pathology of pneumothorax. At the highest level resolution, 0.63 lp/mm, the radiologists' diagnoses were categorized as: "complete agreement," "minor disagreement" and "moderate disagreement." The two "moderate disagreement" ratings resulted from the radiologists' inability to decide on the hemothorax diagnosis.

In the Image Diagnosis Tests, and for the first and second resolution levels (0.079 and 0.157 lp/mm), no consistent pathology was diagnosed by the radiologists. At the third level, 0.315 lp/mm, 3 of the 4 radiologists diagnosed pneumothorax with confidences ranging from low to high. At the best television resolution of 0.63 lp/mm, all 4 radiologists provided the correct diagnosis of left lung pneumothorax. The confidence levels in this diagnosis ranged from medium to high.

The results of this case indicate that a resolution of 0.315 lp/mm is satisfactory for diagnosing pneumothorax with a reasonable confidence level.

CASE TEN, SKULL WITH FRACTURE

For this case, the fact that the pineal was not calcified reduced the choices of diagnoses in the Diagnoses Evaluation Test to either a fracture or normal. The first use of the fracture diagnosis was at the second resolution level, 0.315 lp/mm, and was made by 3 of the 4 group 1 radiologists. At the third level, 0.63 lp/mm, all 4 radiologists gave the fracture

diagnosis, with 2 indicating "probably correct" and the other two as "definitely correct." At the fourth level, 1.2 lp/mm, 3 of the 4 radiologists felt that the fracture diagnosis was "definitely correct."

In the Image Diagnosis Test, the results were basically the same as for the Diagnoses Evaluation Test. Thus, at the second level, 0.315 lp/mm, all group 2 radiologists diagnosed a fracture with confidences from low to medium-high. At the third and fourth levels (0.63 lp/mm and 1.2 lp/mm), 3 of the 4 radiologists diagnosed multiple fractures with high confidence levels. The fourth radiologist merely diagnosed parietal skull fracture.

The results for this case indicate that resolution levels greater than 0.315 lp/mm (but not more than 0.63 lp/mm) are necessary to define both the fracture and its extent. The lower resolution would be acceptable if a zoom capability were available to increase twofold the image resolution.

TV-BASED IMAGE RESULTS BY DIAGNOSIS

NORMAL CHEST CASE DIAGNOSES (CASES 2, 4, 5, and 9)

In general, when working at the lowest resolution levels, results were better when the diagnosis was "normal" than when it was one of the other four. At the lowest resolution level, 0.157 lp/mm, the evaluations of the normal fell one each into these categories: "complete agreement," "minor disagreement," "moderate disagreement" and "serious disagreement." The one "serious disagreement" resulted from "cannot decide" evaluations and not from erroneous diagnoses. Thus, the normal diagnosis requires the least resolution.

HEMOTHORAX DIAGNOSES FOR CHEST CASES (CASES 2, 3, 4, and 9)

In general, the diagnosis for hemothorax was accurate at the fourth resolution level (0.63 lp/mm), and not quite as accurate at the third resolution level (0.315 lp/mm). For cases 2 and 3, one and 2 radiologists, respectively, had a "cannot decide" at the third resolution, but at the fourth resolution level all 4 radiologists agreed (in direction) with their readings of the films. For case four, all 4 radiologists agreed in direction with their readings of the film at both the third and fourth resolution levels. Only for case nine were there 2 radiologists who still had "cannot decides" at the fourth resolution level (the other 2 radiologists agreed in direction). In case nine, there was no hemothorax pathology and uncertainty as to what level of hemothorax pathology to look for could have caused anomalous results. The resolution required to define a hemothorax diagnosis is likely to be closer to the 0.63 lp/mm resolution than 0.315 lp/mm.

PNEUMOTHORAX DIAGNOSES FOR CHEST CASES (CASES 2, 4, and 9)

At the second resolution level, 0.157 lp/mm, the Diagnoses Evaluation Test results for the group 1 radiologists for pneumothorax diagnoses had 7 of the 12 evaluations (58 percent) of "cannot decide." One of the remaining 12 was erroneous, relative to the diagnostic evaluation based upon film reading. At the third level, 0.315 lp/mm, 75 percent

(9 of 12) of the diagnostic evaluations were generally correct. The 3 remaining evaluations were "cannot decide." At the highest level, 0.63 lp/mm, all of the diagnosis evaluations agreed almost exactly with the film reading results.

Based upon these test results, the resolution necessary for pneumothorax diagnosis is generally in the range of 0.315 to 0.63 lp/mm. The latter resolution would likely be satisfactory for most cases of diagnosis of pneumothorax.

ATELECTASIS DIAGNOSES FOR CHEST CASES (CASES 2, 3, 4, and 9)

The results by group 1 with atelectasis diagnoses at the first and second levels (0.079 and 0.157 lp/mm) were generally in the "major disagreement" category. These resolutions, while presenting the general visual data of gross atelectasis, did not provide sufficient detail to discern between atelectasis and other similar pathologies. At the third level, 0.315 lp/mm, 3 of 4 radiologists' evaluations were categorized as follows: "complete agreement," (2) "minor disagreement," and "moderate disagreement."

These results indicate that a system resolution equivalent to 0.315 lp/mm could be used for diagnosing atelectasis. This would be particularly true when patient history, symptom data, and a close-up scanning capability were available.

PNEUMONIA DIAGNOSES FOR CHEST CASES (CASES 2, 4, 5, and 9)

For the evaluation of the pneumonia diagnoses at the first resolution level, 0.079 lp/mm, 3 radiologists indicated "cannot decide" evaluations for three of the four cases, as shown in Appendix F. For the third and fourth levels (0.315 lp/mm and 0.63 lp/mm, respectively), the radiologists' diagnostic evaluations were categorized as "in agreement" or "minor disagreement."

From the results of these cases, a system resolution of 0.315 lp/mm is considered feasible for a pneumonia diagnosis and, considering that pneumonia is frequently well localized, this system resolution may be adequate if a close-up scanning capability is available.

DISCUSSION

GENERAL CONSIDERATIONS

In the introductory section of this report, it was noted that the function of the RMDS in the transfer of radiographic information will generally be to support intermittent, emergency diagnostic efforts at remote sites. The RMDS is not expected to be employed for routine radiographic screening for early stages of a pathological condition. Thus, the resolution requirements for the RMDS equipment will not be as stringent as other applications might dictate.

The cases which were used in this study are assumed to be typical of those which might be transferred by the RMDS equipment. With the exception of the pelvis of case six, all cases represented either normal or obviously abnormal pathological conditions. The radiologists were able to separate normal from abnormal conditions with a high degree of accuracy.

Furthermore, even though no symptomatic information, other than the images, was available to the radiologists, they were able to provide a large percentage of reasonably accurate diagnoses.

The testing involved the use of only radiologists in the evaluation of radiographic information. In practice, the RMDS is likely to be used in situations where the alternative will be to allow general practitioners (GPs), MDs, or hospitalmen to evaluate the radiographs and to prescribe treatment. Thus, in assessing the efficiency of a certain degree of resolution for the RMDS, there is a question of whether or not the radiologist can perform better using the video image than can a GP or hospitalman using the original radiograph. Some light can be shed on this question by research conducted by Dr W Wilson of Long Beach Memorial Hospital.* His research used a Westinghouse Slow-Scan system to compare the performance of radiologists using television images against the performance of GPs using original radiographs. This testing showed that, for cases with positive diagnoses, the radiologists had an accuracy 30 percent better than the GPs. Even greater improvement in accuracy can be expected when the performance of radiologists is compared to that of hospitalmen.

It should be noted that the tests reported upon in this document were directed toward the analysis of the RMDS radiographic transmission resolution requirements. Several other RMDS equipment parameters are also important to the visualization process for radiographic data, but were not examined. These include noise and dynamic range.

CAMERA AND TRANSMISSION SYSTEM NOISE

The higher the system noise levels, the more difficult will be the visualization of radiographic data and, thus, the lower the accuracy of any diagnostic process. It is expected that system resolution requirements might be reduced if improvements can be achieved in lowering the noise level.

DYNAMIC RANGE OF CAMERA AND DIGITATION PROCESS

A substantial mismatch exists between the wide dynamic range displayed in radiographs and the limited dynamic range in the imaging capabilities of standard television camera system such as were used in the tests. In addition, limiting the transmission times in digital systems requires limiting the number of bits which are used to represent each picture element. It is certain from other reported research that 6 to 8 bits per picture element will provide satisfactory gray-level resolution. Six bits per picture element is expected to be satisfactory for use in the RMDS. Improvements in the area of dynamic range can also be expected to improve system resolution.

In the tests reported upon herein, the radiologists had neither patient histories and symptom information nor the option to select the area of the radiograph and the concomitant magnification to be used. Furthermore, it should be noted that the testing was conducted by one radiologist at a time under highly artificial conditions. These factors tend to reduce the diagnostic performance that might otherwise have been achieved. The operational conditions under which the RMDS will be used will allow the radiologists, as part of a team effort, to

---

* Private correspondence with Dr W Wilson concerning unpublished research conducted by him while at the University of Nebraska.

scrutinize the RMDS-transferred image, weigh it against patient history and the clinical observations of the on-site medical personnel, and to request other radiographic views or additional test data needed to reach a conclusive diagnosis. Thus, while the testing reported herein may superficially indicate a lower performance level than desired at the resolutions used, the test conditions must be weighed in any final conclusions.

COMMENT ON FILM-READING RESULTS

In general, the results pertaining to film-reading indicate that one cannot expect radiologic film interpretation by individual radiologists, in isolation, to lead to an unequivocally specific diagnosis. The exception to this arises in those cases of obvious symptomatic conditions. In part, this result stems from the fact that the diagnoses provided for evaluation did not indicate the degree of the pathological condition. Thus, a conservative radiologist might hesitiate in unequivocally ruling out a pathological condition merely because the film did not present an obvious case of that pathology.

In many cases, the radiographic presentation was such that more than one pathology might be plausible. Final determination of the diagnosis would then require patient-history information. In some cases, the film reading results were simply erroneous. The erroneous results can be attributed to either inexperience of the radiologists, or, in some cases, to hasty analysis of the film.

A significant conclusion related to the film-reading results is that, due to the ambiguities of these results, no single set of *a priori* diagnoses can be used as a basis for evaluating the television-based test results for different radiologists. Instead, the television-based test results for each radiologist can be compared only reasonably to the radiographic film reading results for that individual radiologist.

RESOLUTION REQUIREMENTS

An examination of table 7 which summarizes the results of the previous section by case and, by diagnosis, confirms assumptions made prior to the testing; different anatomy, pathology, and traumas require different resolutions for diagnosis. A somewhat surprising conclusion from these results given in table 7 is that only relatively low values of resolution are required to diagnose many cases.

Figure 4 shows a summary of the data of table 7. In this figure, the plotted points represent the cumulative sum of the percent of test cases requiring a resolution equal to or less than that shown on the abscissa. A curve is plotted between these points. It should be noted that the points plotted represent the results from 15 discrete test data points. The fitting of a curve to these points should not be construed to mean that extrapolation along this curve is completely accurate. Nevertheless, this curve does illustrate a salient point. The point is that a continuum of resolutions exists which relate to the fraction of cases that can be effectively diagnosed at that resolution. The distribution and types of cases selected for this study were intended to approximate the distribution and types of cases which could be expected in an implementation of the RMDS. Thus, it is presumed here that the data of figure 4 present at least a rought approximation of the relative utility of the various resolutions for an operational RMDS.

TABLE 7. SUMMARY OF RESOLUTION REQUIREMENTS DATA.

Case	Diagnosis	Minimum Acceptable Resolution	Comments
1	Fracture	0.157 lp/mm	
2	Chest/atelectasis	0.315 lp/mm	
3	Chest/massive injuries	>0.63 lp/mm	
4	Chest/normal	0.315 lp/mm	For moderately advanced pathology
5	Chest/active TB	0.315-0.63 lp/mm	0.157 lp/mm would be adequate with 4:1 magnification
6	Pelvis/fracture of right pubic ramus	>1 lp/mm	Magnification not an adequate routine substitute
7	Sinus/sinusitis	0.42 lp/mm	
8	Skull/normal	0.63 lp/mm	For obvious fracture conditions
9	Chest/pneumothorax	0.315 lp/mm	
10	Skull/fracture	0.315-0.63 lp/mm	For obvious fracture conditions
2, 4, 5	Chests/normal diagnoses	>0.157 lp/mm	For obvious pathology
2, 3, 4, 9	Chests/hemothorax diagnoses	0.63 lp/mm	
2, 4, 9	Chests/pneumo-thorax diagnoses	0.315-0.63 lp/mm	
2, 3, 4, 9	Chests/atelectasis diagnoses	0.315 lp/mm	
2, 4, 5, 9	Chests/pneumonia diagnoses	>0.315 lp/mm	

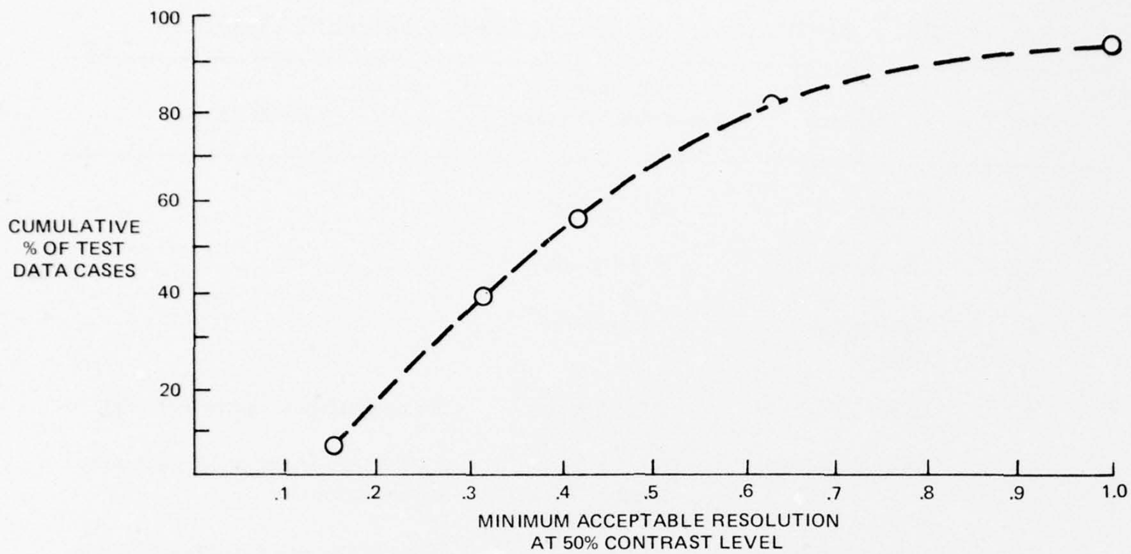

Figure 4. Cumulative fraction of cases for which resolution was acceptable.

Interpretation of the data of figure 4 requires that certain considerations germane to acquiring these data be examined. A significant consideration relative to the data presented is in the meaning of the resolution numbers. An earlier section of this report discussed the recording procedures. It was stated there that a bar pattern was used to set the system resolution. This was accomplished by defocusing the camera lens until the bars for the desired spatial frequency were at about half their highest contrast. It should be noted at this point that resolution cannot be considered a bandpass function, with unit response to the cut-off resolution and zero response thereafter, as shown in figure 5a. Rather, for a defocused lens, the actual system response would resemble the curve shown in figure 5b. Thus, the test recordings were made with a certain spatial frequency, ω_c, at a reduced contrast, r. For frequencies above ω_c, the system still passed information, albeit at reduced contrast. The resolution numbers used in the foregoing portions of this report represent approximately the 50-percent contrast level. Since most system resolutions are specified at 5-to-10 percent levels, the test results herein presented must be adjusted to increased resolution values to be compatible with the normally given resolution numbers.

Another consideration, relative to the resolution data presented, is that the given resolution numbers represent those for imaging the full field of the radiographs used. In some cases (eg, chest films with tuberculosis), a somewhat lower resolution than the minimum acceptable values presented could be used. This would be possible if the system had a zoom capability and could image a smaller portion of the radiograph at increased image resolution but with the same overall system resolution. Thus, a system with a resolution of 0.3 lp/mm over a 30-by-30 cm field could be designed to provide a 0.9 lp/mm resolution when zoomed to a 10-by-10 cm field. The system resolution constraint then would be the maximum resolution required for the largest field to be viewed. The feasibility of this approach largely rests on consideration of cases such as number 6, the pelvis with a fracture of the right pubic ramus. In this case it was determined that a resolution of over 1 lp/mm was required. In

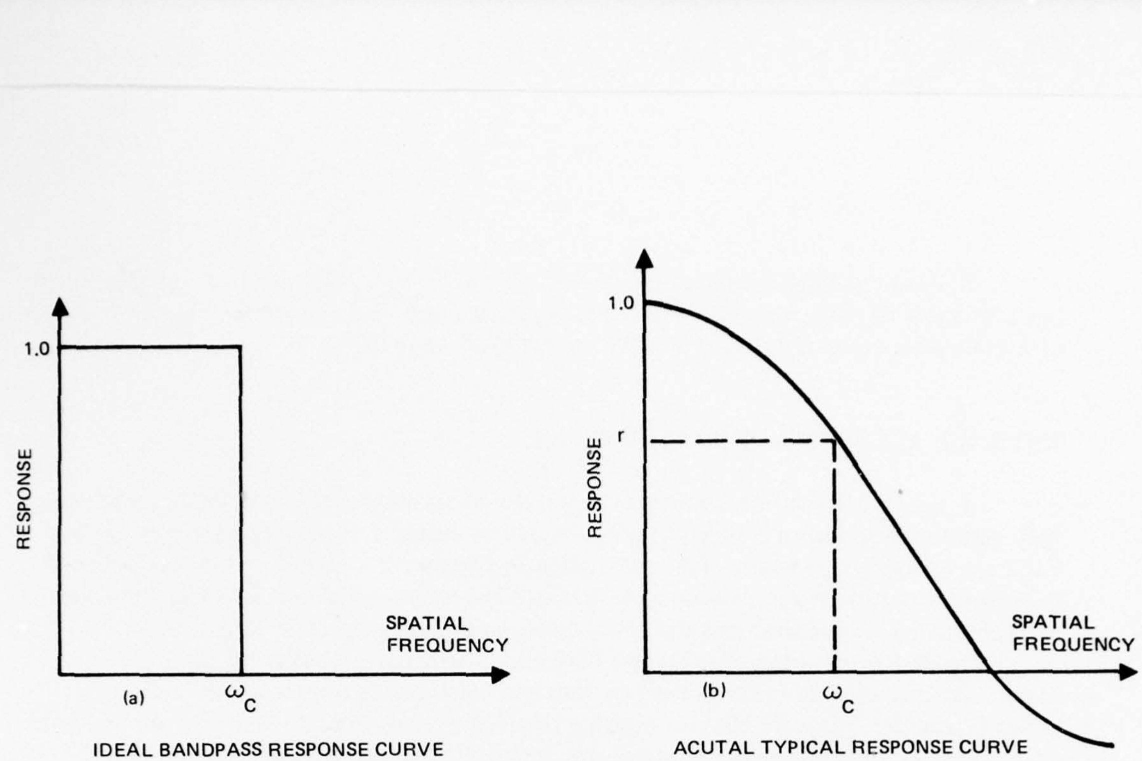

Figure 5. Ideal versus actual response curves for defocused lens.

order to effectively use a zoom capability, one needs to know the area on which to concentrate. In case six, even at the best resolution, the Image Diagnosis Test showed a marginal capability to define the diagnosis. (Note that this was an unprompted diagnostic situation.) In the Diagnoses Evaluation Test, the best television resolution (1.0 lp/mm) proved adequate, but the radiologists had been prompted as to where to look. Thus, the concept of decreasing the system resolution and relying upon a zoom capability depends upon having sufficient information to know upon what area to concentrate. This information can come either from viewing the whole film at reduced resolution (at the film) or from prompting based upon knowledge of patient history and symptoms.

As stated in previous portions of this report, the radiologists were provided with no patient information other than that of the test image. This is undoubtedly an abnormal condition for evaluation of radiographic images. Based upon the several cases examined in this study, appropriate patient histories and symptom information would very likely have allowed the same accuracy of diagnosis at resolutions lower than those quoted in table 7. Further, if the test images were concurrently being evaluated by more than one radiologist, providing a joint diagnosis, it is anticipated that the resolution requirements for diagnosis would be less stringent than as indicated in table 7.

A further consideration relative to the results of these tests is that the participating radiologists were introduced to the tests rather abruptly and with no prior preparation. Thus, the radiologists, as far as is known, did not know about the nature of the tests until asked to participate. They had never tried to diagnose from television images (other than using fluoroscopy, which is an entirely different situation). It seems reasonable to expect that, with some preparation for, or training in diagnosing from television images, the participating radiologists could have performed much better in the tests. This could have resulted in reduced resolution requirements over those presented in table 6.

Another consideration is that the diagnoses presented for evaluation in the Diagnoses Evaluation Test may, in some cases, have been ambiguous. For example, when asked to evaluate a fracture diagnosis, even with the best resolution, the radiologist might not feel certain that he had not overlooked some extremely fine fracture line. Similar ambiguities existed in evaluation of pneumonia and tuberculosis in the chest-film cases.

In both the Diagnoses Evaluation Test and the Image Diagnosis Test, it would have been desirable to have more information about why a radiologist made the choices he did, and upon what visual symptoms or assumptions these were based.

OVERALL ACCURACY OF PERFORMANCE

A significant consideration in defining the performance of radiologists in analyzing radiographic information is their overall accuracy in terms of normal-abnormal evaluation. Each radiographic case in this test can be unambiguously classified as normal or abnormal. It is useful to examine the frequency with which the radiologists defined normal radiographs as abnormal (false-positive) and abnormal radiographs as normal (false-negative).

Table 8 summarizes the false-positive and false-negative results for each of the test cases. These results are broken down by film and video image results and by radiologist group (1 and 2). Thus, for the film-reading results and each case, the number of radiologists evaluating that case as normal, or abnormal, is listed in the appropriate column. Similar information is provided on the video-image results. In addition, for the video-image results, the resolution at which the consensus shown was reached is given.

A significant result from these data is that, with the exception of some false-positive readings for case eight, the radiologists were always able to separate normal and abnormal conditions at one of the video-image resolutions used. Further, there were no overall false-negative readings, and in all the cases, except case eight, the separation of normal from abnormal cases occurred at less than the best video image resolution used.

SUMMARY AND CONCLUSIONS

SUMMARY

Preliminary tests were conducted using a video recording and television system to determine a quantitative value for the resolution required for diagnosis of radiographic images. The tests were of 2 types: (1) the radiologists were asked to evaluate diagnoses presented to them, based upon what they saw in the test image (Diagnoses Evaluation Test); and, (2) the radiologists provided unprompted diagnoses, based upon what they saw in the test images (Image Diagnosis Tests).

Eight radiologists participated in the testing and reviewed 50 images representing 10 test cases. For each test case, 4 television images of differing resolution and the original radiograph were evaluated.

The Diagnoses Evaluation Test results were evaluated, using a qualitative rating system for comparing the responses of the radiologists for the television images to their responses when viewing the original film. A majority-vote analysis technique was used to assess what performance might be achieved with radiologists consulting with each other on their film readings.

By considering the detailed and analyzed results of the Diagnoses Evaluation Test and weighing these results against the Image Diagnosis Test results, conclusions were reached on the minimum resolution requirements for each test case. Additionally, for the 5 diagnoses common to the 5 chest cases, an analysis of diagnostic performance was conducted. This analysis yielded data on minimum resolution requirements based upon diagnosis type.

TABLE 8. OVERALL FALSE-POSITIVE AND FALSE-NEGATIVE RESULTS.

CASE NUMBER	ACTUAL CONDITION	FILM RESULTS(1)				VIDEO IMAGE RESULTS(1)				
		NORMAL(1)		ABNORMAL(2)		NORMAL(2)		ABNORMAL(2)		
		I	II	I	II	I	II	I	II	RESOLUTION(3) (lp/mm)
1	Abnormal			4	4			4	4	0.079(4)
2	Abnormal			4	4			4	4	All
3	Abnormal			4	4			4	4	All
4	Normal	4	4			4	4			0.315
5	Abnormal			4	4			4	4	0.315
6	Abnormal			4	4			4	4	0.63
7	Abnormal			4	4			4	4	0.315
8	Normal	2	4	2		2	2(5)	2	1(5)	1.2
9	Abnormal			4	4			4	4	0.315
10	Abnormal			4	4			4	4	0.63

(1) Shaded areas indicate results agreeing with actual condition

(2) Number of responses in each group

(3) Lowest resolution value to achieve this response set

(4) Lowest resolution used to reach conclusions shown

(5) One radiologist in this group evaluated a "Cannot Decide"

TEST PROTOCOL CONCLUSIONS

The test as designed and conducted contained some shortcomings. Some of these shortcomings were:

(a) Some of the diagnoses presented for evaluation were ambiguous in that they didn't specify the degree of the pathology for which to search;

(b) No interactions with the individually tested radiologists were used to determine the basis for their diagnoses or evaluations;

(c) The radiologists were not able to review the images in consultation with each other;

(d) No preliminary preparations were conducted to optimize the performance of the radiologists in the tests; and

(e) The lack of any patient history of symptomatic data allowed ambiguous results in the film-reading analyses.

In any future tests of this nature, it is recommended that the above factors be corrected. Future tests should be multiphased. In the first phase, the radiologist should be familiarized with, or trained in, reading degraded televised radiographs. A second phase should involve a comparison of individual and teamed radiologists, evaluating various resolution images with no patient history data. In the final phase, test cases would be presented with patient history and symptoms. It is further recommended that, if possible, the tests be conducted with a digital system to assess the effects of image quantization on diagnosis.

RESOLUTION CONCLUSIONS

The resolution results presented here are based fully upon very limited tests and resulting test data. As such, any adaptation of these results should be done with circumspection.

An adequate nominal RMDS system level resolution for transmitting radiographic data for diagnosis, when supplemented with patient history and symptom information, is 0.5 line pair per millimeter at the film under the following conditions:

(a) A 50-percent contrast level of the system modulation transfer function; and

(b) An imaged field size of about 30 by 40 cm. This size is adequate for viewing a chest from the top of the clavicle to the costiophenic angles and is more than needed for the width of the chest.

This resolution is equivalent to approximately 300 television lines per picture height at 50-percent contrast.

APPENDIX A

PHOTOGRAPHS OF THE RADIOGRAPHS USED IN THE TESTS

This appendix contains photographs of the 10 radiographs used in the tests of this report. The photographs show the radiographs which were presented to the radiologists, except that patient information has been omitted for this publication.

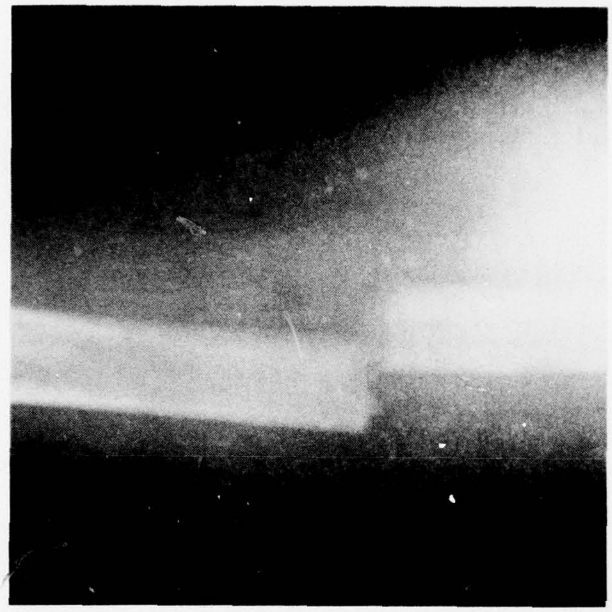

Figure A1. Case one, bone fracture.

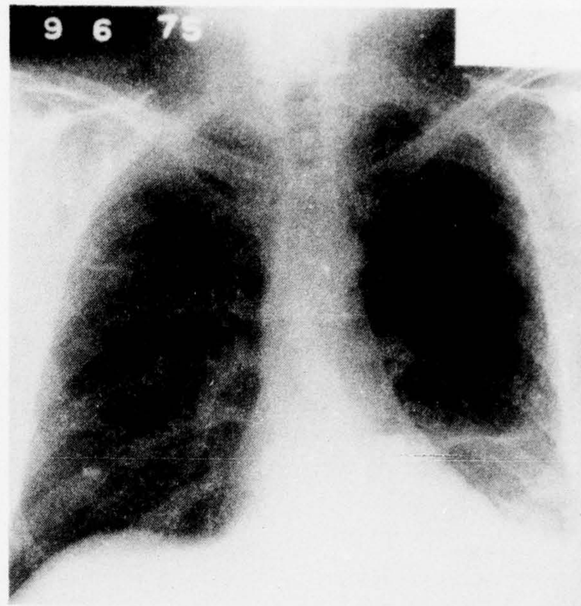

Figure A2. Case two, chest with collapsed lower left lobe.

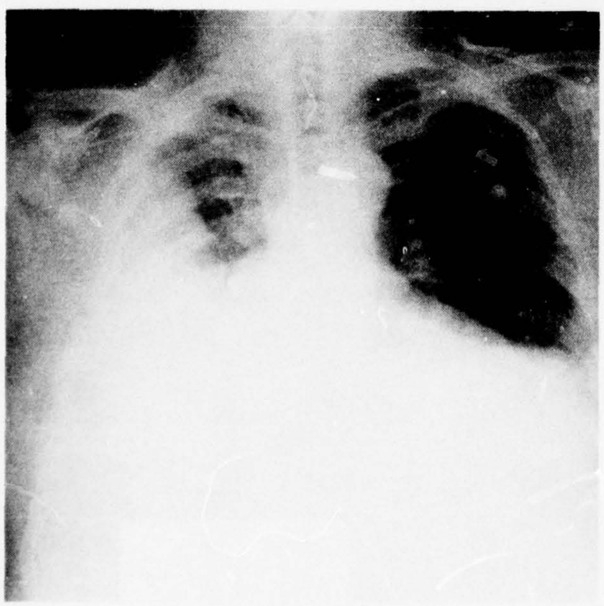

Figure A3. Case three, chest with multiple injuries.

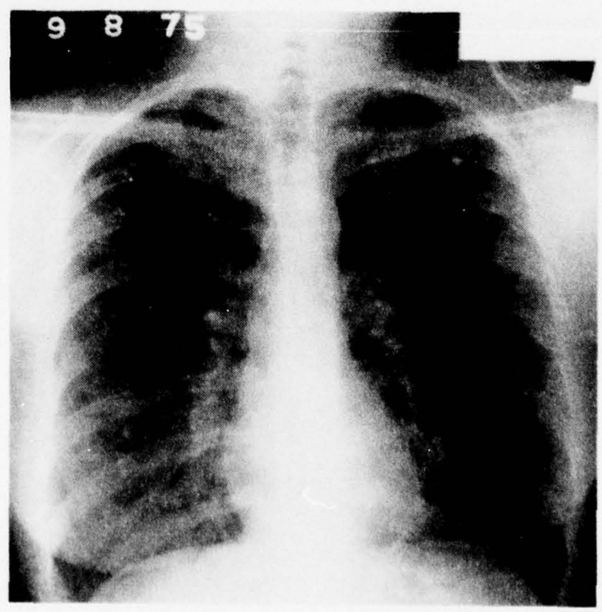

Figure A4. Case four, normal female chest.

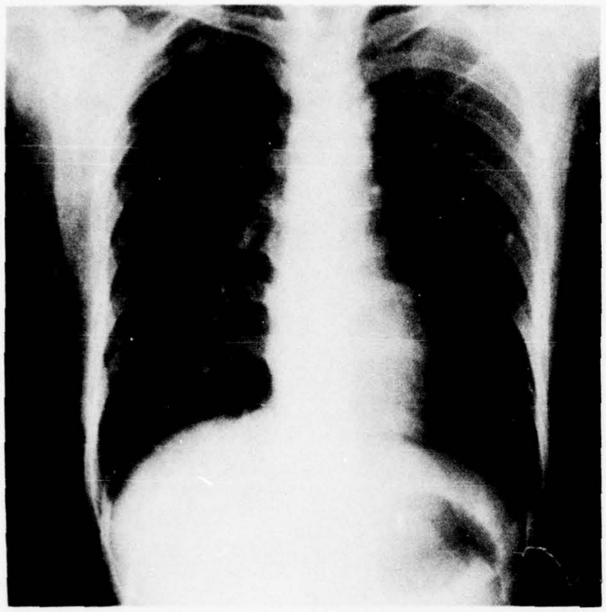

Figure A5. Case five, chest with active tuberculosis.

Figure A6. Case six, pelvis with fracture of right pubic rami.

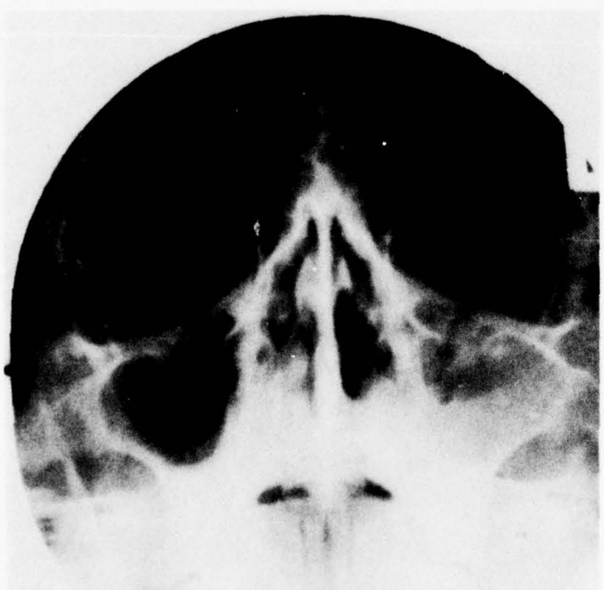

Figure A7. Case seven, sinus with sinusitis.

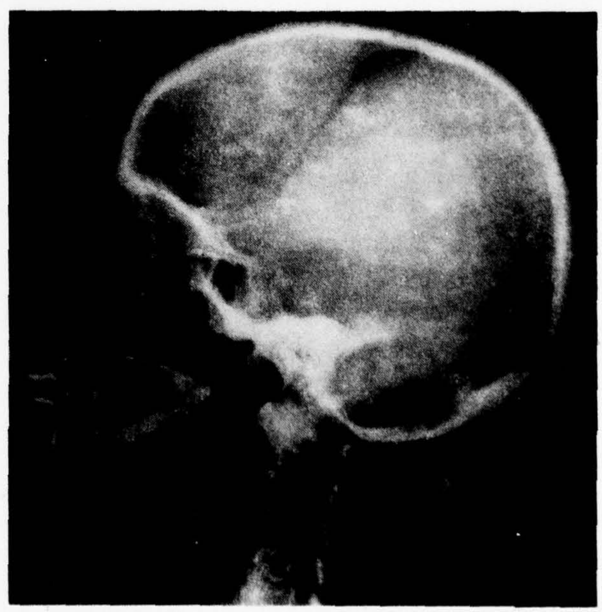

Figure A8. Case eight, normal lateral skull.

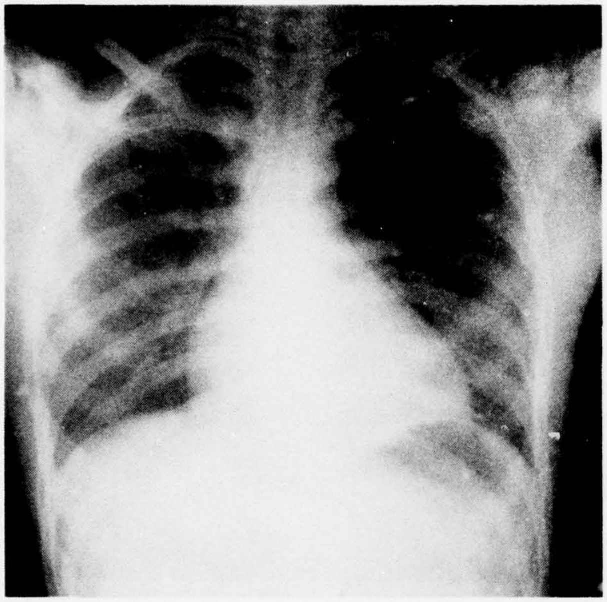

Figure A9. Case nine, chest with pneumothorax (at expiration).

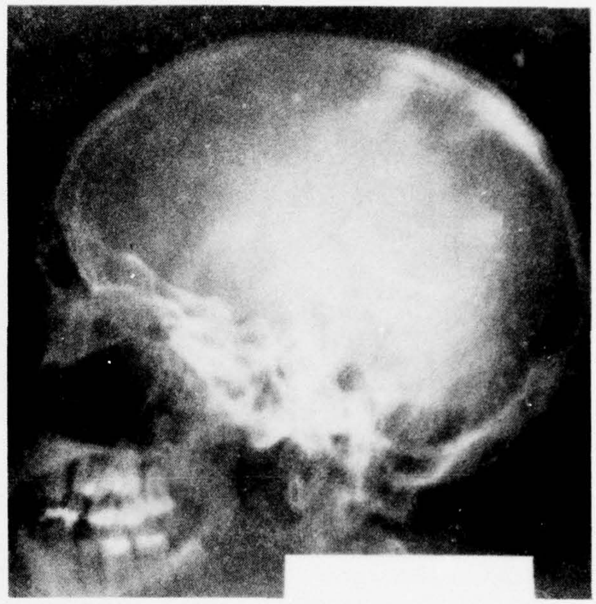

Figure A10. Case ten, skull (lateral) with multiple injuries.

APPENDIX B

PHOTOGRAPHS OF THE RECORDED TELEVISION IMAGES USED IN THE TESTS

This appendix contains illustrations of the television images used in the tests of the report. The images in this appendix are presented in the order they were shown to the radiologist. For each subject, the resolution progresses from poorest (a) to best (d). These photographs were made from polaroid prints of a television monitor. As such, they contribute significant degradation beyond the degraded television images.

The resolution values quoted in the figure captions are those measured at the radiograph surface.

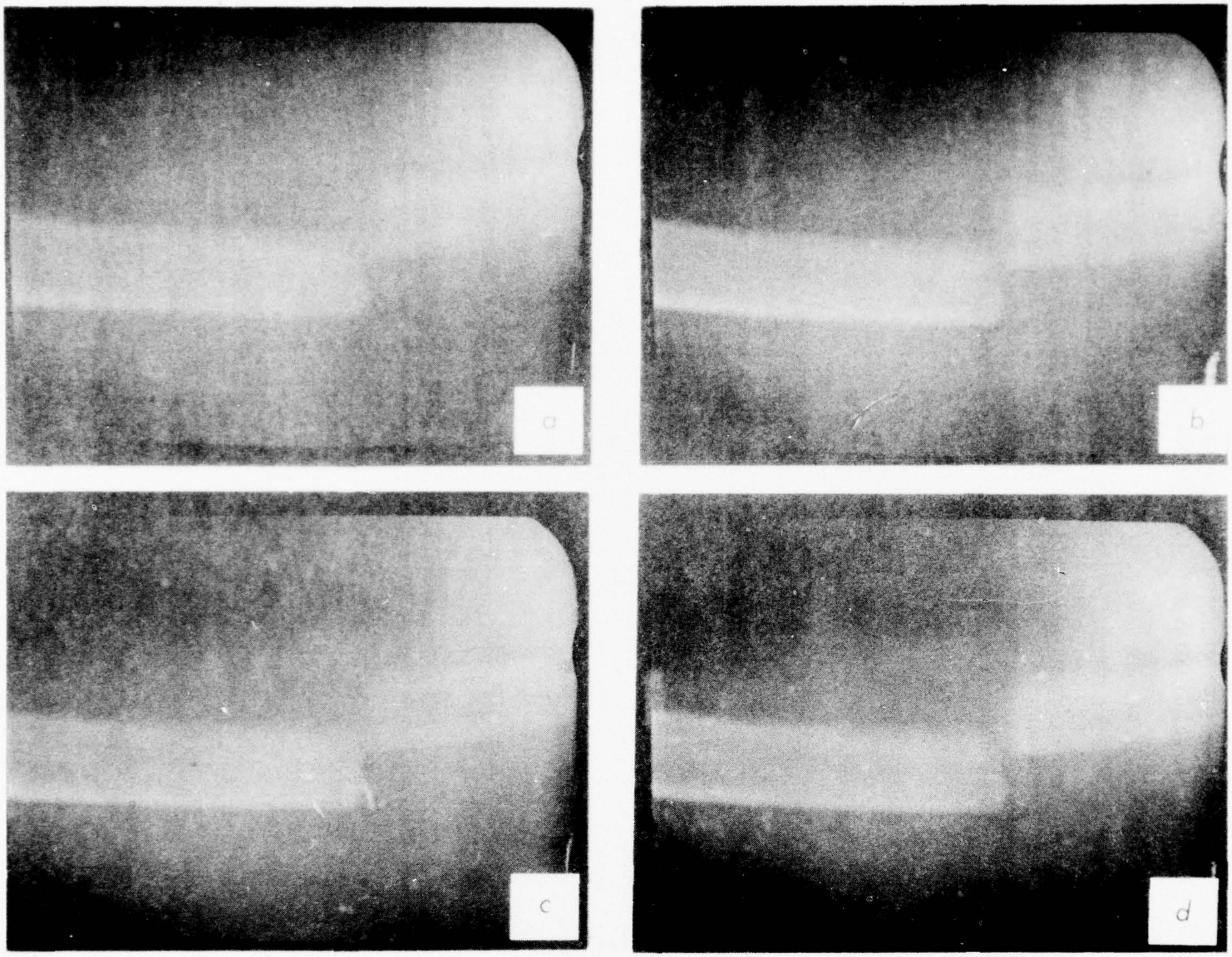

Figure B1. Case one, bone fracture (resolution: a, 0.157 lp/mm, 66 tv lines; b, 0.315 lp/mm, 132 tv lines; c, 0.63 lp/mm, 264 tv lines; and d, 1.0 lp/mm, 419 tv lines).

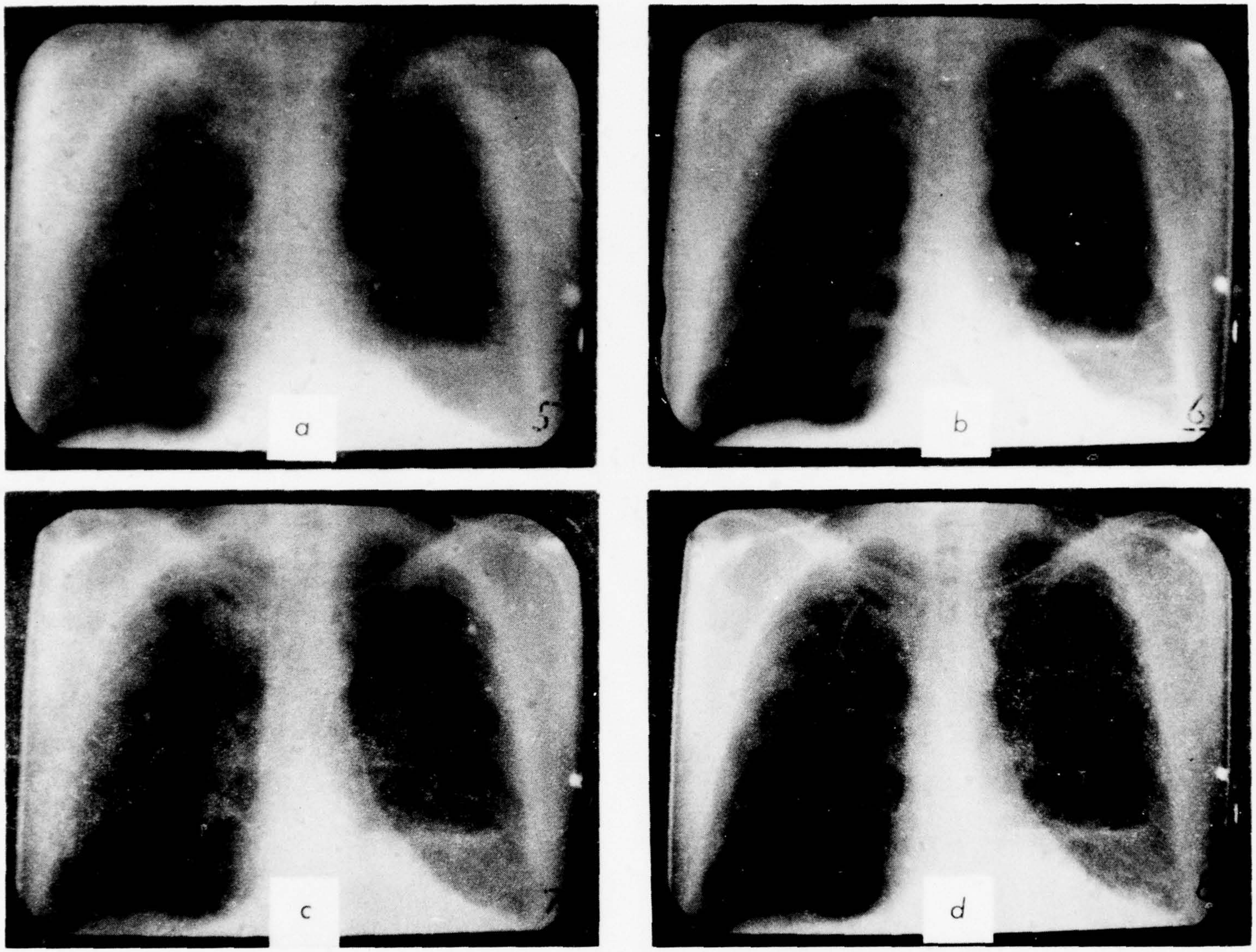

Figure B2. Case two, chest with collapsed left lower lobe (resolution: a, 0.079 lp/mm, 43 tv lines; b, 0.157 lp/mm, 85 tv lines; c, 0.315 lp/mm, 171 tv lines; d, 0.63 lp/mm, 342 tv lines).

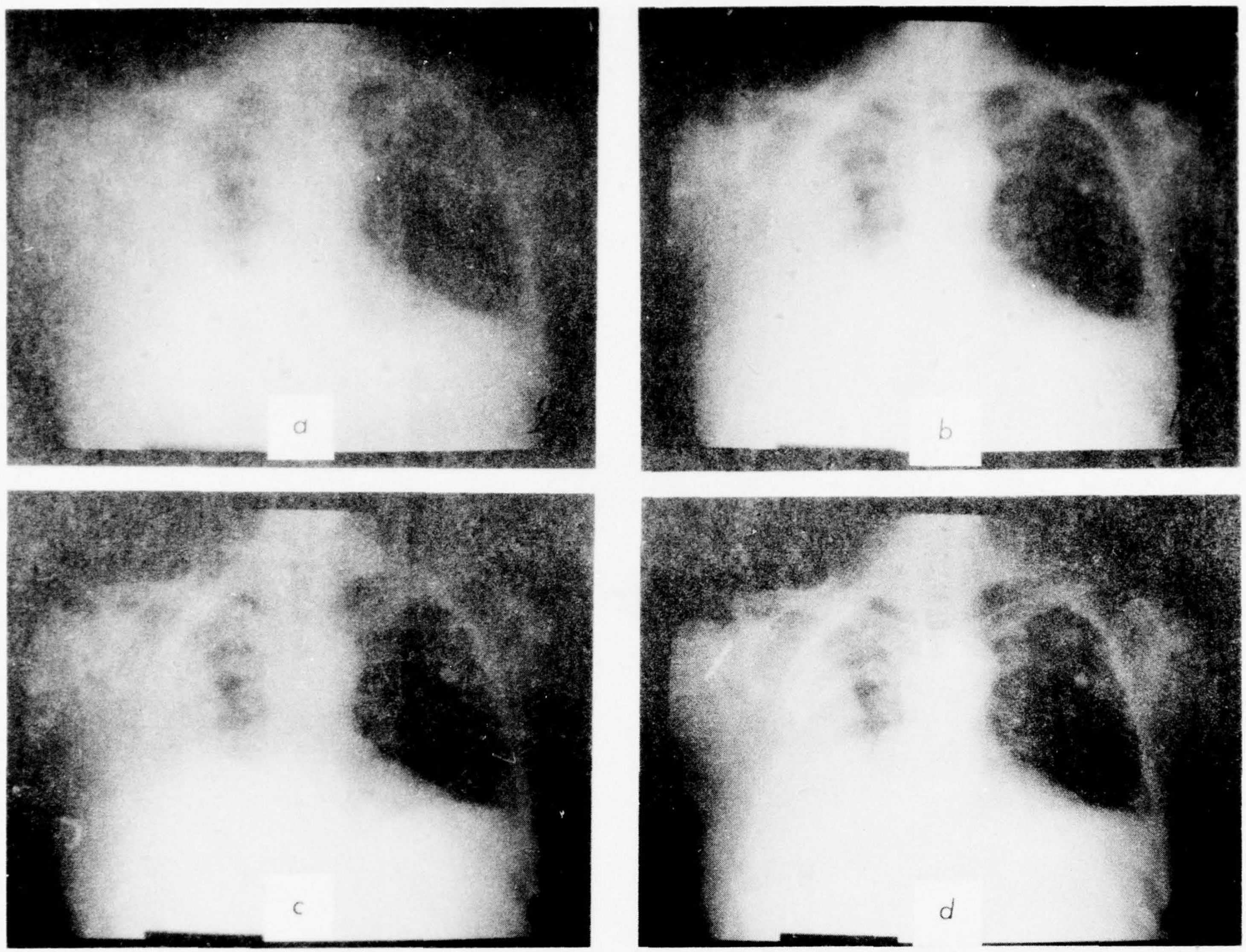

Figure B3. Case three, chest with multiple injuries (resolution: a, 0.079 lp/mm, 50 tv lines; b, 0.157 lp/mm, 99 tv lines; c, 0.315 lp/mm, 198 tv lines; d, 0.63 lp/mm, 396 tv lines).

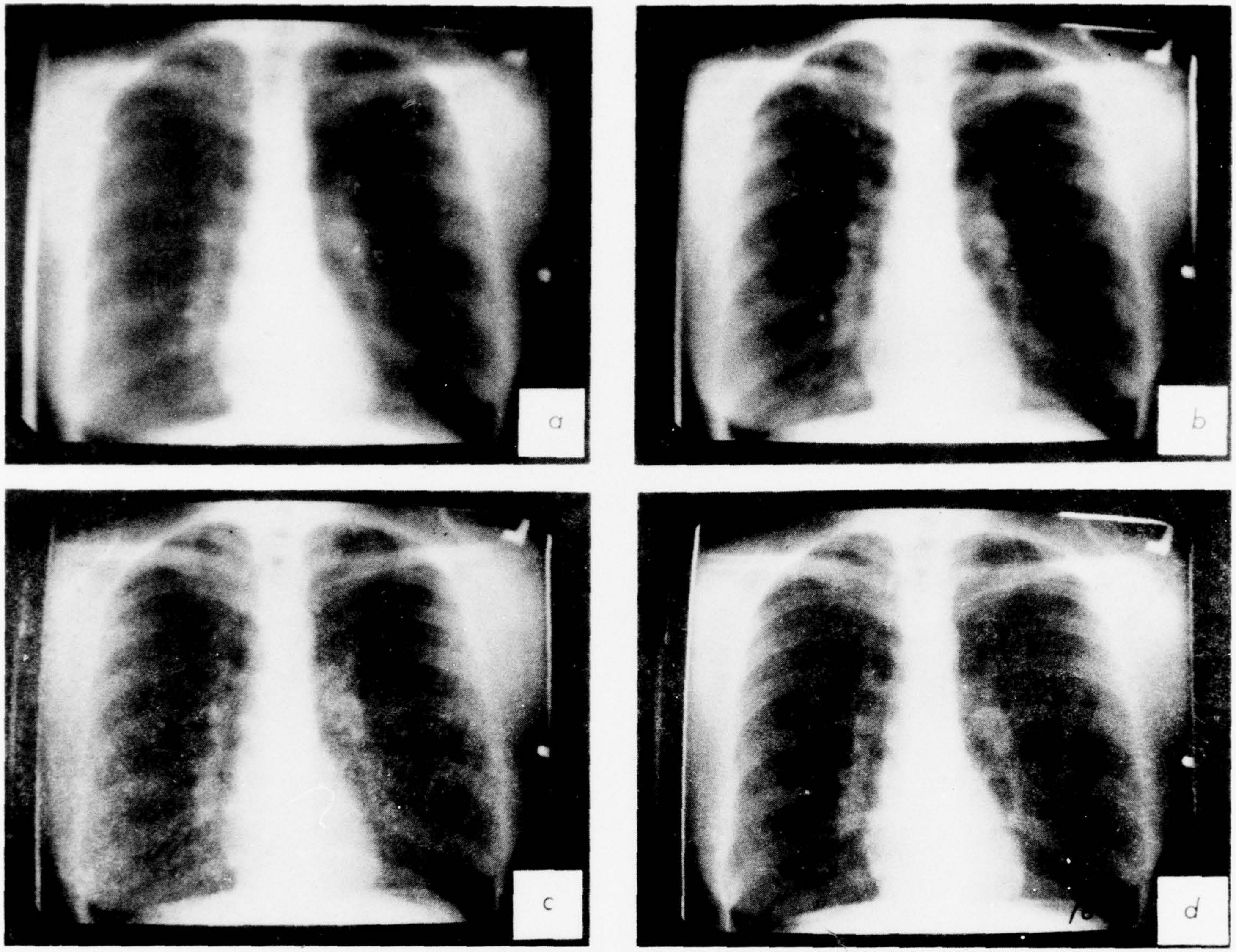

Figure B4. Case four, normal female chest (resolution: a, 0.079 lp/mm, 47 tv lines; b, 0.157 lp/mm, 94 tv lines; c, 0.315 lp/mm, 189 tv lines; d, 0.63 lp/mm, 378 tv lines).

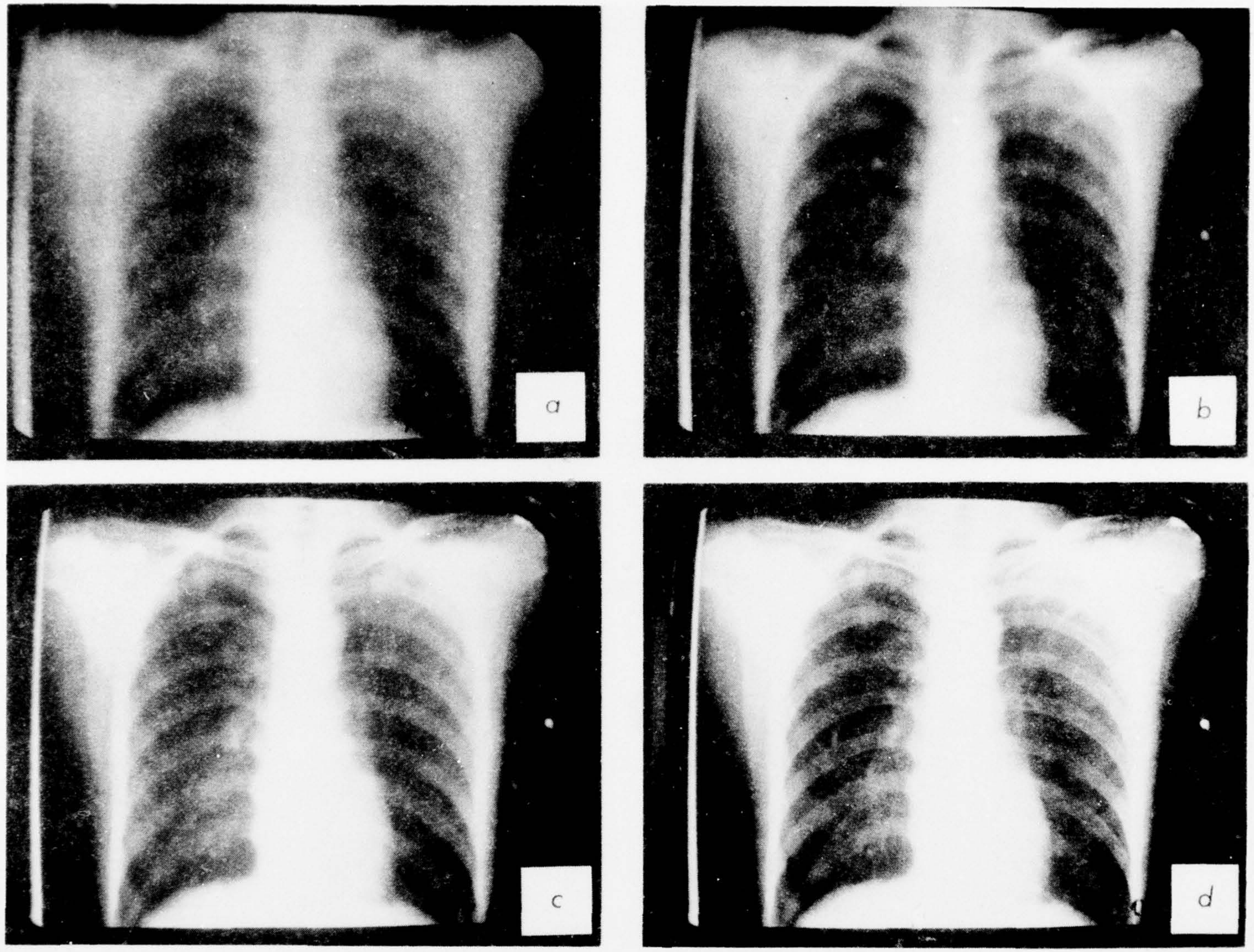

Figure B5. Case five, chest with active tuberculosis (resolution: a, 0.079 lp/mm, 47 tv lines; b, 0.157 lp/mm, 94 tv lines; c, 0.315 lp/mm, 189 tv lines; d, 0.63 lp/mm, 378 tv lines).

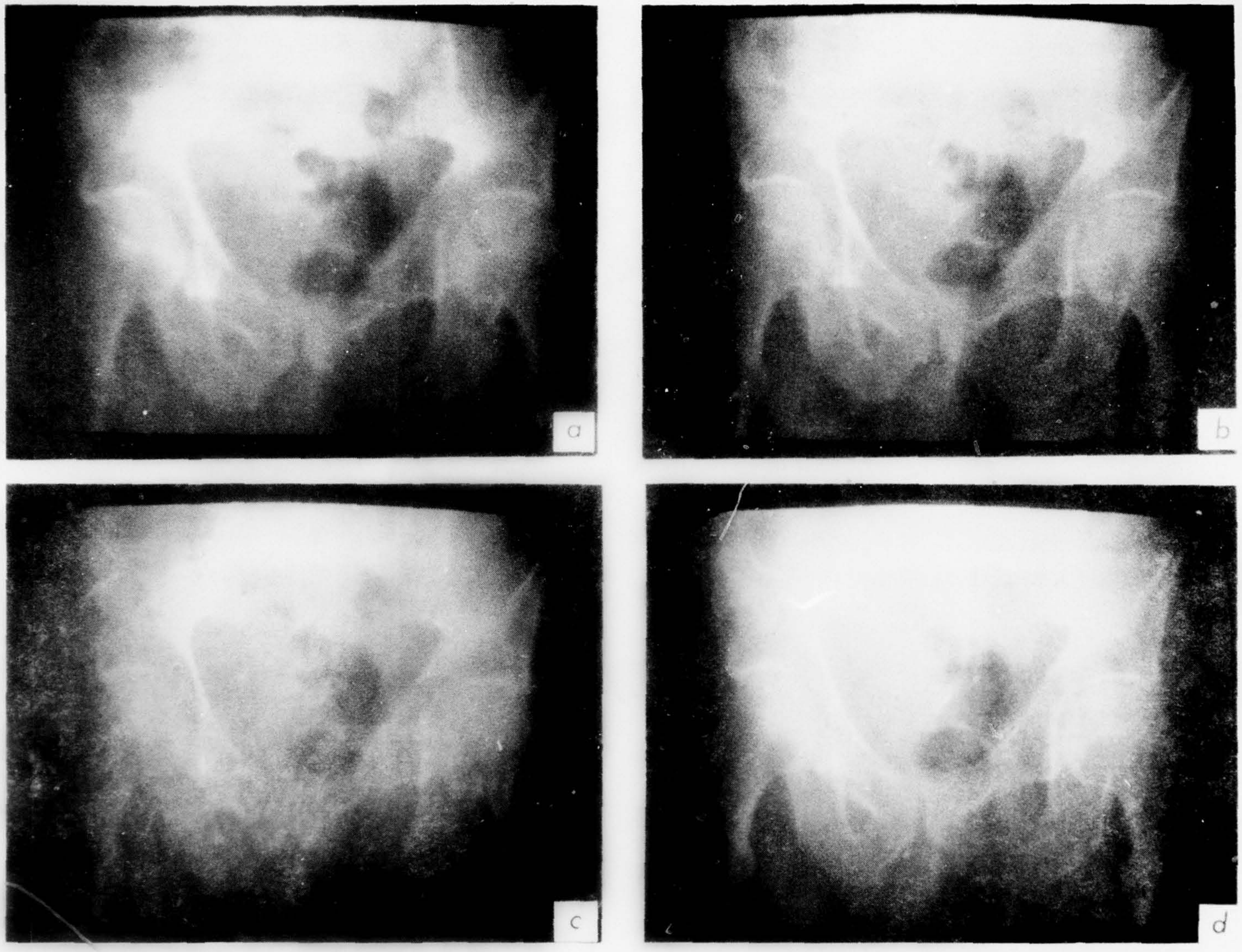

Figure B6. Case six, pelvis with fracture of right pubic rami (resolution: a, 0.157 lp/mm, 80 tv lines; b, 0.315 lp/mm, 160 tv lines; c, 0.63 lp/mm, 321 tv lines; d, 1.0 lp/mm, 510 tv lines).

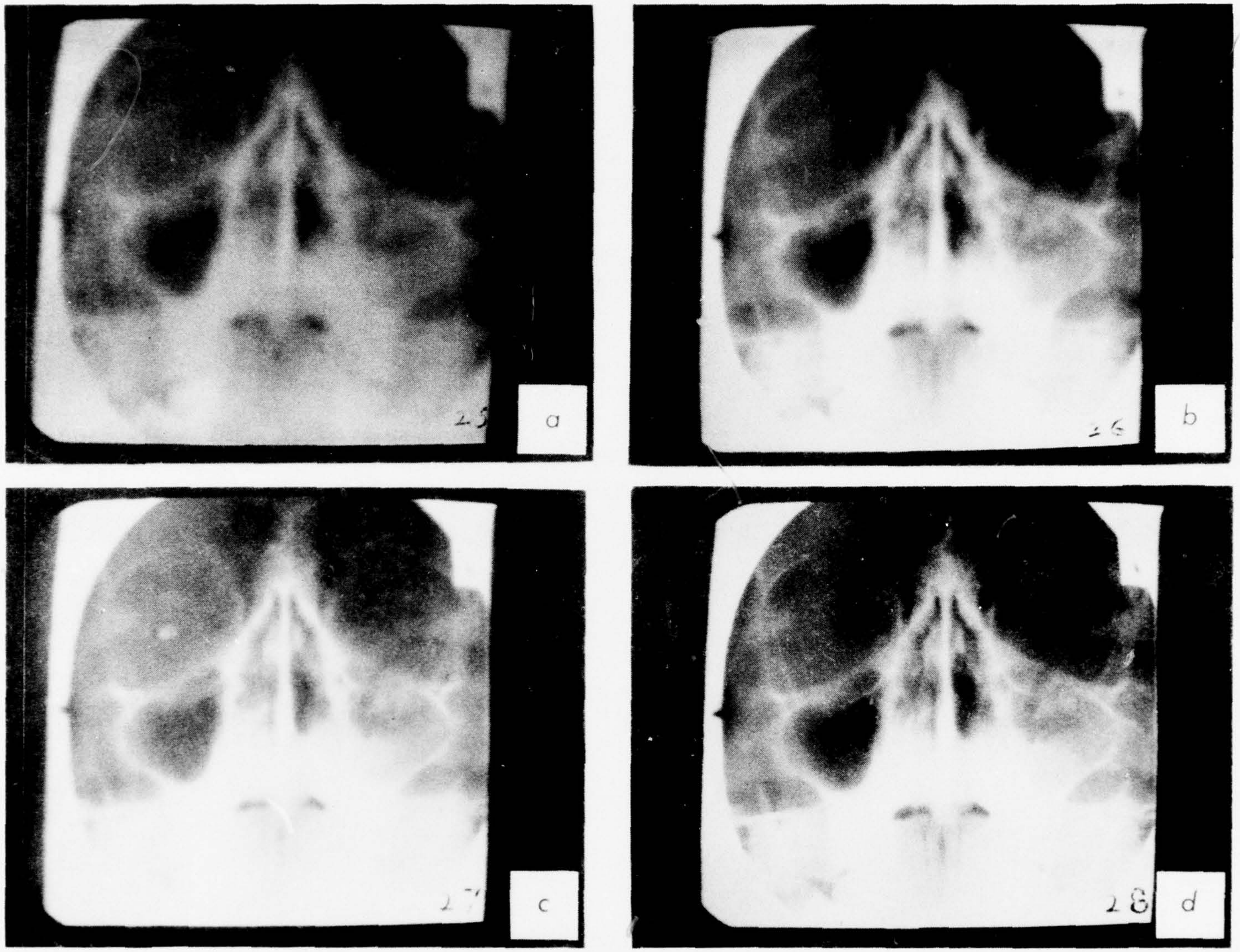

Figure B7. Case seven, sinus with sinusitis (resolution: a, 0.315 lp/mm, 100 tv lines; b, 0.42 lp/mm, 134 tv lines; c, 0.63 lp/mm, 20 tv lines, d, 1.2 lp/mm, 303 tv lines).

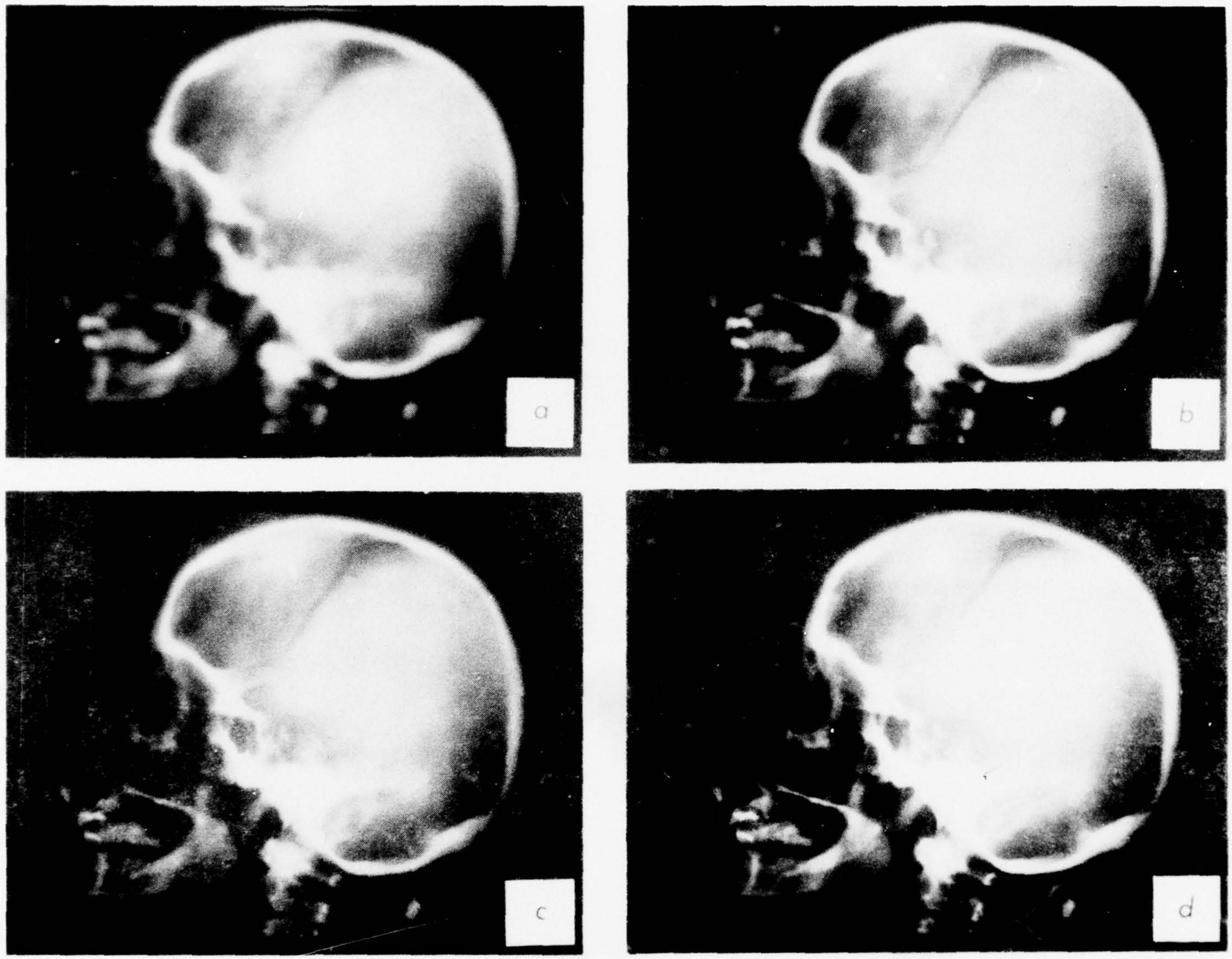

Figure B8. Case eight, normal lateral skull (resolution: a, 0.157 lp/mm, 70 tv lines; b, 0.315 lp/mm, 141 tv lines; c, 0.63 lp/mm, 282 tv lines; d, 1.2 lp/mm, 537 tv lines).

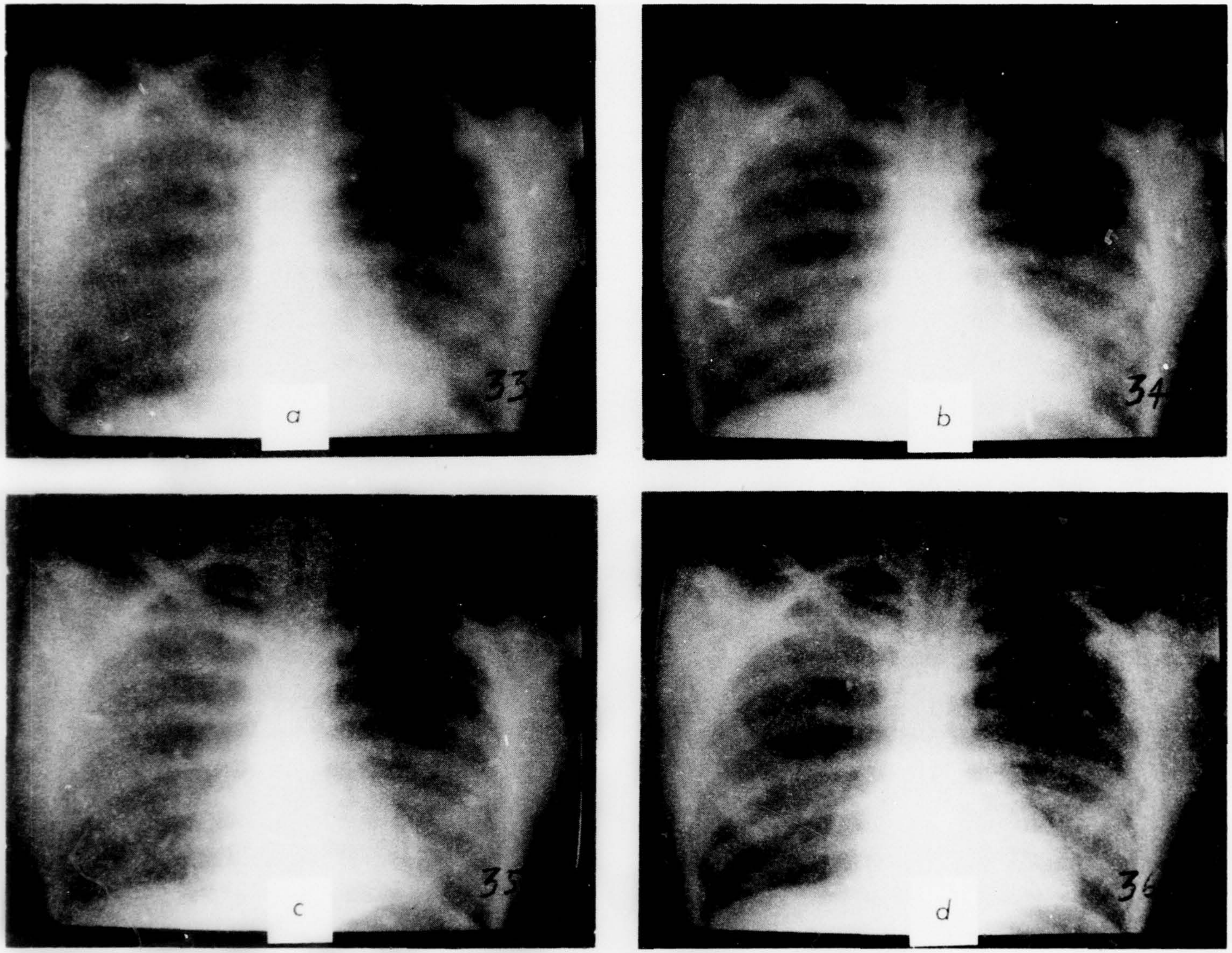

Figure B9. Case nine, chest with pneumothorax (at expiration) (resolution: a, 0.079 lp/mm, 41 tv lines; b, 0.157 lp/mm, 82 tv lines; c, 0.315 lp/mm, 164 tv lines; d, 0.63 lp/mm, 327 tv lines).

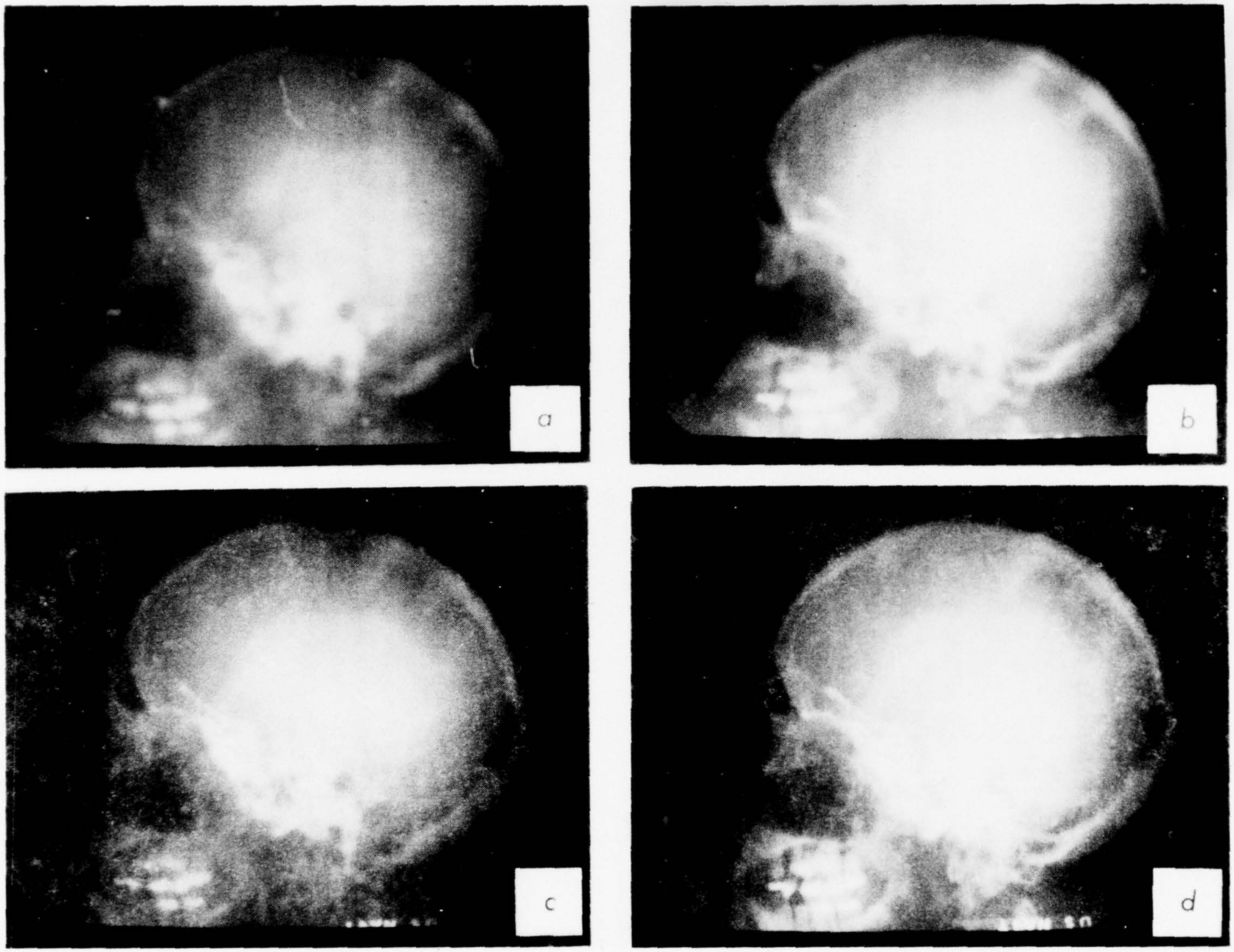

Figure B10. Case ten, skull (lateral) with multiple fractures (resolution: a, 0.157 lp/mm, 70 tv lines; b, 0.315 lp/mm, 141 tv lines; c, 0.63 lp/mm, 282 tv lines; d, 1.2 lp/mm, 537 tv lines).

APPENDIX C

INSTRUCTIONS TO PARTICIPANTS IN DIAGNOSES EVALUATION TEST (GROUP 1)

TEST OF RESOLUTION REQUIREMENTS FOR RADIOGRAPH DIAGNOSIS

BACKGROUND

The Naval Ocean Systems Center, NOSC (formerly the Naval Electronics Laboratory Center), Biomedical Engineering Branch, is currently engaged in a program with the objective of providing improved health care to remotely located Navy operational sites having less than optimum medical staffing. This program envisions the development of the Remote Medical Diagnosis System (RMDS). The RMDS will use hardware equipment to link remote sites, having less than optimum medical staffing, to a diagnosis center where a more complete medical staffing exists. This link will allow for the transfer of medical data from the remote site to the diagnostic center for evaluation and recommendation of treatment. Typical data that will be transferred back to the diagnostic center consist of ECGs, audio data, including voice and heart lung sounds, and video images. These video images will be comprised of not only patient images, but also of radiographic images.

It is anticipated that the data sent from the remote sites to the diagnostic center will be analyzed at the diagnostic center. The diagnosis and any resulting treatment recommendations will be transferred back to the remote site via a voice circuit which will be incorporated as part of the Remote Medical Diagnosis System.

The Remote Medical Diagnosis System will utilize standard Navy communications equipment available at the time of its implementation. Due to the limitations of the communications equipment expected to be available at the time of the RMDS implementation, this system will only be able to transfer still images. Even under these circumstances, though, the transfer of high-resolution images, typical of film-radiography, could involve an inordinately large amount of time. In order to help reduce the amount of time required to transfer a radiograph, it will be necessary to limit the resolution of the images transferred over the RMDS. Therefore, this test in which you are involved is an attempt to resolve the question of what is the minimum resolution required to provide an image from which an accurate diagnosis can be made in a high percentage of the cases.

PROCEDURES

This test will take approximately one hour to complete. During this test you will be asked to view a series of radiographic images on a television monitor with its input originating from a video tape recorder. Each image will be displayed for approximately 90 seconds. For each image presented, you are asked to fill out a test sheet. Each test sheet contains a number of diagnoses relating to that test image. Based upon what you see in the test image on the monitor, you are asked to assess *each* diagnosis as to one of the following classifications:

- Definitely wrong,
- Probably wrong,

- Cannot decide,
- Probably correct, and
- Definitely correct.

Assess each image presented independently of the other images and each diagnosis independently of any other diagnosis. The illustration on the next page shows an example of a filled-out test sheet. Notice that more than one diagnosis can be correct. If you have any comments, reservations, or unresolvable confusion, make appropriate notes in "COMMENTS" at the top of the form. Please print all comments. *Be sure and evaluate each diagnosis for each image.*

For each test image, you are free to adjust the contrast and brightness controls on the monitor to achieve the image representation that you desire.

If for any image you have had insufficient time to complete your test sheet or desire to look at the image longer, please request that the test monitor rerun that image for you. At the completion of the video tape presentation, you will be asked to evaluate the original radiographs. For each radiograph, you will fill out a test sheet in the same manner as was done for the televised images.

Prior to the appearance of each test image on the screen, a number will appear. Check to ensure that the number that appears is the same as the test image number on the top of the sheet which you are about to fill in. "X" the box next to the test image number to indicate that the image number and test sheet numbers agree. In the case of the radiographs, numbered gummed labels will be attached to the radiograph.

CONCLUSION

This experiment will employ a very limited number of radiologists for the test. As such, the results obtained from each radiologist will weigh heavily upon the final results. Thus, it is important that you do not rush the test and that the analysis of each diagnosis be made as accurately as you feel you are able to make it. Evaluate each image presented independently of other images presented and each diagnosis independently of the other diagnoses.

Should you have any questions at this time concerning these instructions or how the experiment will be run, ask the monitor to clarify the situation for you.

Once you feel you understand these instructions and the experiment procedure, please turn to the next sheet and fill in the information requested. When you have completed that sheet, indicate to the experiment monitor that you are ready to begin the testing.

TEST IMAGE # 12 ☑

IMAGE SUBJECT: CHEST

COMMENTS:

#1

Diagnosis:	Definitely Wrong	Probably Wrong	Cannot Decide	Probably Correct	Definitely Correct
ATELECTASIS			✓		

#2

Diagnosis:	Definitely Wrong	Probably Wrong	Cannot Decide	Probably Correct	Definitely Correct
FRACTURE OF SCAPULA	✓				

#3

Diagnosis:	Definitely Wrong	Probably Wrong	Cannot Decide	Probably Correct	Definitely Correct
HEMOTHORAX		✓			

#4

Diagnosis:	Definitely Wrong	Probably Wrong	Cannot Decide	Probably Correct	Definitely Correct
HEMATOMA				✓	

#5

Diagnosis:	Definitely Wrong	Probably Wrong	Cannot Decide	Probably Correct	Definitely Correct
RIB FRACTURE	✓				

APPENDIX D

INSTRUCTIONS TO PARTICIPANTS IN IMAGE DIAGNOSIS TEST (GROUP 2)

TEST OF RESOLUTION REQUIREMENTS FOR RADIOGRAPH DIAGNOSIS

BACKGROUND

The Naval Ocean Systems Center (formerly the Naval Electronics Laboratory Center), Biomedical Engineering Branch, is currently engaged in a program with the objective of providing improved health care to remotely located Navy operational sites having less than optimum medical staffing. This program envisions the development of the Remote Medical Diagnosis System (RMDS). The RMDS will use hardware equipment to link remote sites, having less than optimum medical staffing, to a diagnosis center where a more optimum medical staffing exists. This link will allow for the transfer of medical data from the remote site to the diagnostic center for evaluation and recommendation of treatment. Typical data that will be transferred back to the diagnostic center consists of ECGs, audio data, including voice and heart lung sounds, and video images. These video images will be comprised of not only patient images, but also of radiographic images.

It is anticipated that the data sent from the remote sites to the diagnostic center will be analyzed at the diagnostic center. The diagnosis and any resulting treatment recommendations will be transferred back to the remote site via a voice circuit which will be incorporated as part of the Remote Medical Diagnosis System.

The Remote Medical Diagnosis System will utilize standard Navy communications equipment available at the time of its implementation. Due to the limitations of the communications equipment expected to be available at the time of the RMDS implementation, this system will only be able to transfer still images. Even under these circumstances, though, the transfer of high resolution images, typical of film-radiography, could involve an inordinately large amount of time. In order to help reduce the amount of time required to transfer a radiograph, it will be necessary to limit the resolution of the images transferred over the RMDS. Therefore, this test in which you are involved is an attempt to resolve the question of what is the minimum resolution required to provide an image from which an accurate diagnosis can be made in a high percentage of the cases.

PROCEDURES

This test will take approximately one hour to complete. During this test you will be asked to view a series of radiographic images on a television monitor with its input originating from a video tape recorder. Each image will be displayed for approximately 90 seconds. For each image presented, you are asked to fill out a test sheet. Each test sheet contains space for a number of diagnoses relating to that test image. Based upon what you see in the test image on the monitor, you are asked to provide a list of potential diagnoses. (Please print all entries.) At least one diagnosis should be given. As many diagnoses should be provided as seems justified by what you see in the image. Each diagnosis you provide should be classified

as to your confidence in that diagnosis based upon what you see in the test image. Classify your confidence in each diagnosis by checking one of the following on the test sheet:

- Low,
- Medium-Low,
- Medium,
- Medium-High, and
- High.

Assess each image presented independently of the other images and each diagnosis independently of any other diagnosis you provide. The illustration on the next page shows an example of a filled-out test sheet. If you have any comments, reservations, or unresolvable confusion, make appropriate notes in "COMMENTS" at the top of the form. Please print all comments. *Be sure and evaluate your confidence in each diagnosis for each image.*

For each test image, you are free to adjust the contrast and brightness controls on the monitor to achieve the image representation that you desire.

If for any image you have had insufficient time to complete your test sheet or desire to look at the image longer, please request that the test monitor rerun that image for you. At the completion of the video tape presentation, you will be asked to evaluate the original radiographs. For each radiograph, you will fill out a test sheet in the same manner as was done for the televised images.

Prior to the appearance of each test image on the screen, a number will appear. Check to ensure that the number that appears is the same as the test image number on the top of the sheet which you are about to fill in. "X" the box next to the test image number to indicate that the image number and test sheet numbers agree. In the case of the radiographs, numbered gummed labels will be attached to the radiograph.

CONCLUSION

This experiment will employ a very limited number of radiologists for the test. As such, the results obtained from each radiologist will weight heavily upon the final results. Thus, it is important that you do not rush the test and that each diagnosis be made as accurately as you feel you are able to make it. Evaluate each image presented independently of other images presented and each diagnosis independently of the other diagnoses.

Should you have any questions at this time concerning these instructions or how the experiment will be run, ask the monitor to clarify the situation for you.

Once you feel you understand these instructions and the experiment procedure, please turn to the next sheet and fill in the information requested. When you have completed that sheet, indicate to the experiment monitor that you are ready to begin the testing.

TEST IMAGE # 19 ☑ ✓

IMAGE SUBJECT: CHEST

COMMENTS:

#1

Diagnosis:	CONFIDENCE				
Fracture of scapula	LOW	MEDIUM-LOW	MEDIUM	MEDIUM-HIGH	HIGH
					✓

#2

Diagnosis:	CONFIDENCE				
Rib fracture	LOW	MEDIUM-LOW	MEDIUM	MEDIUM-HIGH	HIGH
				✓	

#3

Diagnosis:	CONFIDENCE				
Hematoma	LOW	MEDIUM-LOW	MEDIUM	MEDIUM-HIGH	HIGH
		✓			

#4

Diagnosis:	CONFIDENCE				
	LOW	MEDIUM-LOW	MEDIUM	MEDIUM-HIGH	HIGH

APPENDIX E

RESOLUTION VALUES FOR TEST IMAGES
(LINE PAIRS/mm AND EQUIVALENT TV LINES)

CASE		LEVEL			
		1	2	3	4
1	BONE FRACTURE	0.157/66	0.315/132	0.63/264	1.0/419
2	CHEST WITH ATELECTASIS	0.079/43	0.157/85	0.315/171	0.63/342
3	CHEST WITH MULTIPLE INJURIES	0.079/50	0.157/99	0.315/198	0.63/396
4	CHEST – NORMAL	0.079/47	0.157/94	0.315/189	0.63/348
5	CHEST WITH ACTIVE TB	0.079/47	0.157/94	0.315/189	0.63/378
6	PELVIS WITH FRACTURE	0.157/80	0.315/160	0.63/321	1.0/510
7	SINUS WITH SINUSITIS	0.315/100	0.42/134	0.63/201	1.2/383
8	SKULL (LATERAL): NORMAL	0.157/70	0.315/141	0.63/282	1.2/537
9	CHEST WITH PNEUMOTHORAX	0.079/41	0.157/82	0.315/164	0.63/327
10	SKULL (LATERAL): WITH FRACTURE	0.157/70	0.315/141	0.63/282	1.2/537

APPENDIX F
DIAGNOSIS EVALUATION TEST DATA

Case / Diagnosis	Resolution Level Evaluation / Radiologist				
	1	2	3	4	Radiograph
1 Bone Fracture					
a. Fracture	++ ABCD	++ ABCD	++ ABCD	++ ABCD	++ ABCD
b. Normal	-- ABCD	-- ABCD	-- ABCD	-- ABCD	-- ABCD
2 Chest with Atelectasis					
a. Atelectasis	+ B 0 AD - C	++ BC + AD	++ BC + AD	++ ABC + D	++ ABC + D
b. Hemothorax	+ C 0 ABD	0 B - ACD	0 B -- A	+ B -- A	+ B -- A
c. Normal	-- ABCD	-- ABCD	-- ABCD	-- ABCD	-- ABCD
d. Pneumonia	+ C 0 ABD	+ BC 0 AD	++ BC + A - D	+ ABC - D	+ ABC - D
e. Pneumothorax	0 BD - AC	0 AD - B -- C	0 D - AB -- C	0 D - B -- AC	0 D - B -- AC
3 Chest with Multiple Injuries					
a. Atelectasis	+ B 0 D - A -- C	++ B 0 AD -- C	++ B + A 0 D -- C	++ B + A 0 D -- C	++ B + A 0 D -- C
b. Fracture of Scapula	0 ABCD	+ BD 0 AC	+ ABD 0 C	++ ABCD	++ ABCD
c. Hemothorax	++ C + B 0 AD	++ BC + D - A	++ BC 0 AD	++ BC + AD	++ BC + AD
d. Hematoma	0 BCD - A	0 BCD - A	+ B 0 ACD	+ ABD 0 C	+ ABD 0 C
e. Rib Fracture(s)	+ C 0 ABD	++ BC + D 0 A	++ BC + A 0 D	++ ABC 0 D	++ ABC 0 D

APPENDIX F (Continued)

Case / Diagnosis	Resolution Level Evaluation / Radiologist 1	2	3	4	Radiograph
4 Chest: Normal					
a. Atelectasis	0 ABD - C	0 BD - AC	- ACD -- B	- D -- ABC	-- ABCD
b. Hemothorax	0 BD - AC	0 B - ACD	- CD -- AB	- D -- ABC	-- ABCD
c. Normal	+ C 0 ABD	+ ABC 0 D	+ ACD - B	++ AC + BD	++ BD + AC
d. Pneumonia	0 ABD - C	0 ABD - C	+ B - ACD	- ABD -- C	-- ABCD
e. Pneumothorax	0 ABD - C	0 ABD - C	0 BD - AC	- AD -- BC	-- ABCD
5 Chest With Active Tuberculosis					
a. Active Tuberculosis	+ C 0 ABD	+ ABC 0 D	+ ABC 0 D	+ ABC 0 D	+ ABC 0 D
b. Bacterial Pneumonia	+ C 0 ABD	+ C 0 ABD	+ C 0 BD -- A	+ C 0 D - AB	+ C 0 BD - A
c. Normal	0 AD - C -- B	0 D - A -- BC	-- ABCD	-- ABCD	-- ABCD
d. Pleural Plaque	0 ABCD	0 ABD - C	+ B 0 AD -- C	+ B 0 D - A -- C	0 BD - A -- C
6 Pelvis With Fracture					
a. Fracture of Left Pubic Ramus	+ B 0 AC - D	0 ABCD	+ C - ABD	+ C 0 B - AD	- BC -- AD
b. Fracture of Right Pubic Ramus	0 BC - AD	+ B 0 ACD	++ B + AD 0 C	++ B + AD - C	++ ABCD

APPENDIX F (Continued)

Case / Diagnosis	1	Resolution Level Evaluation / Radiologist 2	3	4	Radiograph
c. Inter-Trochan Fracture of Left Femoral Neck	0 ACD – B	+ B 0 ACD	+ C 0 AD -- B	+ C – AD -- B	++ C -- ABD
d. Normal	0 ACD – B	0 CD – A -- B	– ACD -- B	– ACD -- B	-- ABCD
7 <u>Sinus with Sinusitis</u>					
a. Facial Fracture	+ C 0 ABD	+ ACD 0 B	+ ACD 0 B	+ ACD – B	– AC -- BD
b. Normal	– A -- BCD	– A -- BCD	-- ABCD	-- ABCD	-- ABCD
c. Sinusitis	+ AB 0 CD	+ B 0 AC – D	+ AB 0 C – D	+ AB 0 C – D	++ B + ACD
d. Soft Tissue Swelling	+ B 0 ACD	+ AB 0 CD	+ BC 0 AD	+ AB 0 CD	++ B + A – C -- D
8 <u>Skull (Lateral): Normal</u>					
a. Fracture	0 ABCD	++ C + B 0 D – A	++ BC – AD	++ C + B – D -- A	++ A – BC -- D
b. Normal	0 ABCD	+ A 0 D – B -- C	+ AD -- BC	+ AD 0 B -- C	++D + C – B -- A
c. Pineal Displacement	0 ABCD	+ C 0 AB – D	+ C 0 B – AD	+ C – ABD	0 ABC – D

APPENDIX F (Continued)

Case / Diagnosis	Resolution Level Evaluation / Radiologist 1	2	3	4	Radiograph
9 <u>Chest With Pneumothorax</u>					
a. Atelectasis	0 ABD – C	0 D – ABC	+ B 0 C – D –– A	++ C + B – D –– A	++ C + B –– AD
b. Hemothorax	+ BC 0 D – A	+ BC 0 D – A	0 BC – D –– A	0 BC – D –– A	–– ABCD
c. Normal	0 AD –– BC	0 AD – B –– C	– D –– ABC	– BD –– AC	–– ABCD
d. Pneumonia	+ C 0 ABD	+ AC 0 BD	0 C – BD –– A	– BD –– AC	–– ABCD
e. Pneumothorax	0 ABD – C	0 BD – AC	++ AB + CD	++ ABC + D	++ ABCD
10 <u>Skull (Lateral) With Fracture</u>					
a. Fracture	0 ABCD	+ ABC 0 D	++ AB + CD	++ ABC + D	++ ABCD
b. Normal	0 ABCD	0 D – AC – B	– D –– ABC	– D –– ABC	–– ABCD
c. Pineal Displacement	0 ABCD	+ C 0 ABD	0 ABCD	+ A 0 BCD	0 ABCD

Legend:

–– Definitely Wrong	A – Radiologist A
– Probably Wrong	B – Radiologist B
0 Cannot Decide	C – Radiologist C
+ Probably Correct	D – Radiologist D
++ Definitely Correct	

APPENDIX G

SAMPLE OF RADIOLOGIST COMMENTS FOR DIAGNOSIS EVALUATION TEST

Test Image	Case	Resolution (lp/mm)	COMMENTS (Verbatim)
12	3	0.63	"Cannot see any rib fractures, if any exist."
16	4	0.63	"Will be very difficult to show small pneumothorax, even at best resolution here."
19	5	0.315	"Can say film is abnormal, but still too gross to tell nature of process."
20	5	0.63	"Can see abnormality fairly well, but can't decide between the various possibilities. However, having a better resolution film would not be likely to answer the question, since these diseases can look like each other."
24	6	1.0	"Even at best resolution, cannot see internal bone detail well, so cannot see small fracture lines."
28	7	1.26	"Very difficult to see soft tissue changes even at best resolution."
32	8	1.26	"Very difficult to see small fracture lines in the skull. Would need fairly obvious fracture."
36	9	0.63	"Still not completely comfortable with this resolution for diagnosis of pneumothorax."
38	10	0.315	"*Possibly* these lines represent fractures, but too gross to do anything except equivocate."
40	10	1.2	"Probably a fracture, but would need other *views* for further determination."
43	3	Radiograph	"Gross rib fractures on original film which did not transmit."*
45	5	Radiograph	"Resolution *not* the limiting factor: even with good resolution, it is not possible to make a specific diagnosis."
46	6	Radiograph	"Fat stripe simulating intertrochanteric fracture cannot definitely be differentiated by transmitted image."
47	7	Radiograph	"Looked like fracture on transmission, but need original films to say it is not. So, the poor resolution can cause diagnosis of abnormalities which are not real."
49	9	Radiograph	"There is also a pneumomediastinuma which did not transmit* at all."

* Radiologist means that rib fracture did not show up on TV images.

APPENDIX H

CATEGORIZATION OF DIAGNOSIS EVALUATIONS

Case		Resolution Level Category / Radiologist 1	2	3	4
2	Chest With Atelectosis*	--- ABCDM	+ C -- B --- ADM	- C -- B --- ADM	- BC --- ADM
3	Chest With Massive Injuries	--- ABCDM	-- B --- ACDM	- B --- ACDM	- B -- ADM --- C
4	Normal Chest	- C --- ABDM	- C --- ABDM	- ACM -- DM --- B	+ M - ABCDM
5	Chest With Active Tuberculosis**	- C --- A	- C --- A	+ C -- A	+ AC
6	Pelvis With Fracture	--- ABCDM	--- ABCDM	+ B -- ADM --- C	- ADM -- B --- C
7	Sinus With Sinusitis	-- B --- ACDM	-- B --- ACDM	-- BM --- ACD	- B -- AM --- CD
8	Normal Skull	--- ABCDM	-- BC --- ADM	- D -- BC --- AM	- D -- BC --- AM
9	Chest With Pneumothorax	--- ABCDM	--- ABCDM	+ A - D -- B --- CM	+ A - D -- BC --- M
10	Skull With Fracture	--- ABCDM	- ABCM --- D	+ ABM - CDM	+ ABCM - D

* Radiologist C's film evaluation for the hemothorax diagnosis was "Cannot Decide," and thus diagnosis was not used in these quality assessments.

** Radiologists B and D were not evaluated due to multiple "Cannot Decide" evaluating for final film diagnosis.

Note: Majority Vote (M) can be in two categories due to ties.

Legend: Complete agreement (+), Minor Disagreement (–), Moderate Disagreement (––), Serious Disagreement (–––).

APPENDIX I

SAMPLE OF RADIOLOGIST DIAGNOSES FOR IMAGE DIAGNOSIS TEST

Test Image	Case	Resolution (lp/mm)	DIAGNOSIS (Verbatim)	Confidence Level
1	1	0.157	"Fracture, transverse"	High
2	1	0.315	"Fx, transv."	High
3	1	0.63	"Fracture, transverse"	High
4	1	1.0	"Fracture, transverse"	High
5	2	0.079	"Left pleural effusion"	Low
6	2	0.157	"Left pleural effusion"	Medium-low
7	2	0.315	"Pleural scarring"	Medium
8	2	0.63	"Pleural scar L base"	Medium
9	3	0.079	"Partial opac. R chest"	High
9			"Calicified nodule L lung"	High
9			"Pleural effusion L base"	Low
10	3	0.157	"Partial opacification R lung"	High
10			"Calcified nodule L lung"	Medium
10			" L pleural effusion"	Low
11	3	0.315	"Partial opac. R chest"	High
11			" R infrahilar mass"	Medium-high
11			"Left pleural effusion"	Medium-low
11			"Chest leads"	Medium-high
12	3	0.63	"Partial opac. R chest, possible loculated pleural eff."	High
12			" R infrahilar mass"	Medium-high
12			" L pleural effusion"	Medium-high
12			"Chest leads"	High

APPENDIX I (Continued)

Test Image	Case	Resolution (lp/mm)	DIAGNOSIS (Verbatim)	Confidence Level
13	4	0.079	"Hyperinflation"	Medium-high
13			"Centrally enlarged pulmonary arteries"	Medium-high
14	4	0.157	"Hyperinflation"	Medium-high
14			"Centrally enlarged pulmonary arteries"	Medium-high
14			" R L Lung density"	Low
15	4	0.315	"Hyperinflation"	Medium-high
15			"Central enlarged pulmonary arteries"	Medium-high
15			" R L lung density"	Low
16	4	0.63	"Hyperinflation"	Medium-high
16			"Central enlarged P.A." ("Would like to have seen castophrenic angle on R ")	Medium-high
17	5	0.079	"Hyperinflation"	Medium-high
17			"Old rib fractures on R "	Low
18	5	0.157	"Hyperinflation"	Medium-high
18			" R U L nodular densities or rib fractures"	Medium
19	5	0.315	"Hyperinflation"	Medium-high
19			"Fractures of 4th & 5th ribs *or* overlying nodules"	High
20	5	0.63	"Hyperinflation"	Medium-high
20			"Nodular densities (2) rib fx vs nodules"	High
21	6	0.157	"? lytic lesion L ischium"	Low
22	6	0.315	"? lytic lesion L ischium"	Low
22			"? sclerotic lesion R ilium"	Low
23	6	0.63	"? lytic lesion L ischium"	Low

* Comments provided by radiologist.

APPENDIX I (Continued)

Test Image	Case	Resolution (lp/mm)	DIAGNOSIS (Verbatim)	Confidence Level
23			"? sclerotic lesion R ilium"	Low
24	6	1.0	"? lytic lesion L ischium"	
24			"? sclerotic lesion R ilium"	Medium
25	7	0.315	"Air fluid level L max. sinus"	Medium-high
26	7	0.42	" L max. sinus mucosal thickening"	High
7	7	0.63	" L max sinus mucosal thickening"	High
28	7	1.26	" L max. sinus mucosal thickening"	High
28			"Absence of bone destruction"	Medium-low
29	8	0.157	"? hole in frontal bone"	Medium-low
30	8	0.315	"Hole in frontal bone"	Low
30			"Lucent area top coronal suture"	Medium-low
31	8	0.63	"Enlarged vascular groove"	Medium
32	8	1.26	"Large vascular groove – normal skull"	Medium
33	9	0.079	" R pleural density laterally"	Low
33			"Borderline cardiomegaly"	Medium-high
34	9	0.157	"Pleural calcification bilaterally"	Medium-low
34			"Borderline cardiomegaly"	High
35	9	0.315	" L pneumothorax"	Medium-low
35			"Borderline cardiomegaly"	High
36	9	0.63	" L pneumothorax"	Medium
36			"Borderline cardiomegaly"	High
37	10	0.157	"Normal"	Low
38	10	0.315	"Skull fx"	Low
39	10	0.63	"Skull fx multiple"	High

APPENDIX I (Continued)

Test Image	Case	Resolution (lp/mm)	DIAGNOSIS (Verbatim)	Confidence Level
40	10	1.2	"Multiple skull fractures"	High
41	1	Radiograph	"Transverse fx femur"	High
42	2	Radiograph	"Pleural scar ± or fluid	High
43	3	Radiograph	" R rib fxs"	High
43			" R pleural lesion"	High
43	3	Radiograph	" L pleural fluid"	Medium-high
43			" R infrahilar density (mass lesion)"	Medium
44	4	Radiograph	"Centrally enlarged pulmonary arteries"	High
45	5	Radiograph	"2 nodules RUL "	High
45			"Hyperinflation"	Medium-high
46	6	Radiograph	"Old fracture R ischium"	Medium-high
46			"Old fracture R pubis	Medium
47	7	Radiograph	"Mucosal thickening L max. sinus"	High
48	8	Radiograph	"Normal"	High
49	9	Radiograph	"Pneumothorax L "	High
50	10	Radiograph	"Multiple skull fractures"	High